ajer

THE ALBERTA JOURNAL
EDUCATIONAL RESEAR

VOLUME XXVIII NUMBER 2 JUNE

PUBLISHED BY
THE UNIVERSITY OF ALBERTA • EDMO

The University of Alberta
75th Anniversary
1982-83

THE ALBERTA JOURNAL OF EDUCATIONAL RESEARCH

A quarterly journal devoted to the dissemination, criticism, interpretation and encouragement of all forms of systematic enquiry into education and fields related to or associated with education.

Published quarterly in March, June, September, December by the
Faculty of Education, The University of Alberta

CONSULTING EDITORS

W. R. Andrews
Queen's University

J. Britton
University of London

J. Calam
The University of British Columbia

B. M. Chung
Chungbuk University, Republic of Korea

K. De Clerck
State University of Ghent

R. N. Evans
University of Illinois at Urbana-Champaign

G. Harman
University of Melbourne, Australia

D. B. Harrison
The University of Calgary

N. L. Hersom
University of Saskatchewan

S. Hunka
The University of Alberta

D. A. MacIver
University of New Brunswick

E. Miklos
The University of Alberta

P. Robinson
Atlantic Institute of Education

R. E. Traub
Ontario Institute for Studies in Education

EDITOR: Doyal Nelson

SECRETARY: A. Onishenko

FACULTY PUBLICATIONS COMMITTEE

N.C. Bhattacharya
R.C. Bryce (Chairman)
A.M. Decore
J. Fris
W.W. Laing
D. Nelson
J.B. Parsons
C.H. Preitz
D.J. Searle
L.R. Wilgosh

Editorial policy and the discussion and disposition of manuscripts are the joint responsibility of the Publications Committee. The views expressed and the accuracy of the statements made are the responsibility of the individual authors. The editor is solely responsible for the editorial comments.

AJER gratefully acknowledges support from the Social Sciences and Humanities Research Council of Canada and from the Alberta Advisory Committee for Educational Studies.

AJER is indexed in the *Canadian Education Index, Current Contents/Social and Behavioral Sciences,* and *Social Science Citation Index;* appropriate articles are abstracted in *Educational Administration Abstracts, Psychological Abstracts, Sociology of Education Abstracts, and Language Behavior Abstracts.*

The subscription rate is $10.00 per year; single copies are $3.00 each. Please make cheques payable to *The Alberta Journal of Educational Research*. All back issues are available; rates supplied on request. Claims for undelivered copies must be received within three months of the month of publication.

Address all communications and manuscript submissions to the Editor, *The Alberta Journal of Educational Research,* Faculty of Education, 845 Education South, The University of Alberta, Edmonton, Canada, T6G 2G5.

SECOND CLASS MAIL REGISTRATION NUMBER 1436

ISSN 0002-4805

The Alberta Journal of Educational Research

Volume XXVIII, Number 2 *June, 1982*

CONTENTS

Broken Homes: Impact on Adolescents 95
P. W. Koziey and L. Davies

A Field-Based Evaluation of Sustained Silent Reading (SSR) in Intermediate Grades 100
E. G. Summers and J. V. McClelland

The Balance Effect on the M-Space by Spatial Interaction 113
R. V. McCormick

Teacher Professional Profile: A Personal and Contextual Analysis 122
Y. L. J. Lam

A Comparison Study of Three Late Immersion Programs 135
S. M. Shapson and E. M. Day

The Australian School Principal's Interpersonal World of Work 149
J. Chapman and Q. Willis

An Exploratory Study into Determinants of Successful Reading Remediation 163
M. Beebe and G. Malicky

School Subjects Attitude Scales 175
V. R. Nyberg and S. C. T. Clarke

BOOK REVIEWS

The Characteristics and Needs of Adults in Postsecondary Education, by L. C. Solmon and J. J. Gordon 188
Reviewed by R. Schultz

Intelligence, Learning, and Action, by R. R. Skemp 190
Reviewed by A. T. Olson

FACULTY OF EDUCATION
The University of Alberta

The Alberta Journal of Educational Research *Vol. XXVIII, No. 2, June, 1982, 95-99*

PAUL W. KOZIEY
and
LEIGH DAVIES
The University of Alberta

Broken Homes: Impact on Adolescents

In this study, the effects of different home types on the personal adjustment of adolescents were investigated. A factorial design utilizing two three-level factors and a two-level factor was used. Three home situations were considered under Factor A, home type: (a) broken by separation or divorce; (b) broken through death of one parent; and (c) intact homes. Factor B, age of child at disruption, was considered when studying disruptions occurring at three ages: between 0 and 5, 6 and 12, 13 and over. Factor C was sex. The sample was 54 grade ten students. Six scales of the California Personality Inventory (CPI) were administered: Femininity, Good Impression, Self-Acceptance, Self-Control, Sociability, and Socialization. There were differences between broken home and intact home adolescents on four of the six CPI scales. No clear-cut trend emerged linking age of home disruption with personal adjustment, and no sex differences were found.

Recent years have witnessed an increase in the incidence of marital rupture in North America. Although enumerations are admittedly incomplete and do not take into account the multiplicity of factors causing broken homes, attention is often drawn to the negative psychological impact the divided home has upon young children (Hammond, 1979; Kalter, 1977). Attempts to understand the psychological factors associated with this stressful interpersonal environment have not been satisfactory (Kelly & Wallerstein, 1976; Rubin & Price, 1979).

Some psychologists point to the fact that children from broken homes often experience feelings of inadequacy or rejection, due in part to a tendency to blame themselves for their parents' marital difficulties (Hammond, 1979; Kelly & Wallerstein, 1976; Rubin & Price, 1979). Perhaps the most unfortunate eventual

---

Paul Koziey (PhD, University of Wyoming) is an Associate Professor in Educational Psychology at the University of Alberta. His major teaching, research and clinical interests lie in the areas of Human Potential Psychology and Family Therapy with a particular focus on the utilization of unconscious resources through Neuro Linguistic Programming and Ericksonian Trance.

Leigh Davies (MEd, University of Alberta) is a counseling psychologist in the city of Edmonton with special interests in psychodiagnosis and developmental psychology.

outcome of marriage breakdown is the apparent predisposition of children from broken homes to experience, later in life, difficulties similar to those of their parents. This is reflected by a much higher percentage of divorces and separations occurring in groups from families broken by divorce, separation, or desertion of either parent (Rubin & Price, 1979).

Also of considerable concern to some researchers has been the inadequate role model provided for the child of a broken home. Lack of appropriate sex-models in early childhood has been found to result in delayed or incomplete sex-role identification (Cox & Cox, 1979; Hammond, 1979; Hetherington, Cox & Cox, 1978).

In general, then, the personality characteristics of children from broken homes are those associated with an incompletely socialized person. One could expect children from these homes to exhibit, to a degree greater than the general population, personality characteristics indicative of poor social knowledge and poor adjustment (Cox & Cox, 1979; Kalter, 1977; Wallerstein & Kelly, 1975).

The Study

In the present study, an attempt was made to determine the nature and extent of the psychological impact of different home types on personal adjustment of adolescents. The investigators were primarily interested in the differences between the personality characteristics of adolescents from homes experiencing marriage breakdown and from intact homes; however, a group of adolescents from homes broken by the death of one parent was included in the study for purposes of comparison. Assessment of personality was accomplished by means of selected scales from the California Psychological Inventory (CPI). The scales chosen were Self-Control (Sc), Socialization (So), Sociability (Sy), Self-Acceptance (Sa), Femininity (Fe), and Good Impression (Gi).

A factorial design utilizing two three-level factors and a two-level factor was used, thereby allowing analysis with the ANOVA model which considers main and interaction effects (Winer, 1962). Factor A, home type, comprised three levels: (a) homes broken by separation or divorce; (b) homes broken through death of one parent; and (c) intact homes. Factor B, the age of children at time of disruption of the homes, comprised three levels: (a) disruption occurring between ages 0 and 5; (b) ages 6 and 12; and (c) ages 13 and over. Factor C was sex.

The following hypotheses were formulated and tested in an attempt to isolate personality factors associated with home disruption:

1. Adolescents from homes broken by separation, divorce, or death will show a greater degree of personal maladjustment than adolescents from intact homes.
2. Girls from homes broken by separation, divorce, or death will show a greater degree of personal maladjustment than boys from similar circumstances.
3. The younger the child when the home was broken by separation, divorce, or death, the greater the degree of personal maladjustment.

The sample consisted of 54 students: 18 from homes broken by separation or divorce; 18 from homes broken by the death of one parent; and 18 from intact homes. The first student questionnaire was administered to all grade 10 and special vocational students of a large Catholic high school. The sex of the child, nature of home disruption, and age of child at the time of disruption were secured from this questionnaire. Approximately three weeks later, the total sample was administered the six selected scales of the CPI.

Results

The data were analyzed using a three-way analysis of variance (Winer, 1962) with three subjects per cell. As indicated in Table 1, some significant differences were found among the groups on four of the six CPI scales. No significant differences between broken and intact home groups or between male and female students of the total sample were found on the Sociability (Sy) or Self-Acceptance (Sa) scales. Consequently, these results are not reported.

TABLE 1

THREE-WAY ANALYSIS OF VARIANCE ON FOUR SCALES OF THE CALIFORNIA PERSONALITY INVENTORY

Source	df	Self-Control		Socialization		Femininity		Good Impression	
		MS	F	MS	F	MS	F	MS	F
A (Home)	2	238.907	3.917*	249.388	6.388*	52.241	4.67*	100.722	4.669
B (Age)	2	128.241	2.102	63.388	1.623	4.796	.428	15.555	.698
C (Sex)	1	12.518	.205	.074	.002	332.518	29.728*	96.000	4.449
AB	4	22.685	.372	20.194	.517	10.629	.950	14.111	.654
AC	2	4.018	.066	1.352	.034	17.241	1.541	8.166	.378
BC	2	320.907	5.262*	105.129	2.693	7.018	.627	87.389	4.050
ABC	4	100.074	1.641	99.157	2.540	18.907	1.690	37.556	1.741
ERROR	36	60.981		39.037		11.185		21.574	

$p < .05$

Analysis of the data did not confirm the existence of significant differences in psychological adjustment in adolescents from broken homes and from intact homes on all six scales of the CPI. However, the significant differentiation that did occur on four of the six scales does mirror a trend toward greater maladjustment in adolescents from homes broken by separation or divorce and by death of the father.

No clear-cut trend emerged linking the age of the child when the home was broken with low adjustment on the six CPI scales. In addition, girls from broken homes were not, as a group, significantly more maladjusted than boys from similar circumstances, as measured by the six CPI scales.

Discussion

Broken and intact home groups were found to differ significantly in their responses to the Self-Control (Sc), Socialization (So), Femininity (Fe), and Good Impression (Gi) scales. While this differentiation did not occur to the extent hypothesized, there is an indication of the extent of impact the loss of one parent

can have on the personality on a child from a broken home regardless of whether the home had been broken by separation, divorce, or death.

Significant differences were found to exist between all males and all females of the sample on the Fe and Gi scales. However, no significant interaction effect was discovered between home type and sex. These differences are in all likelihood a result of normal sex-personality variations which are not attributed to the influence of a particular home type.

Significant interaction effects were obtained between the factors of age and sex on the Sc and Gi scales. Subsequent tests of simple main effects on the Sc scale revealed that females experiencing home disruption before age six performed significantly better than males from all home types. The interaction effect between age and sex on the Gi scale was similarly examined. Males experiencing home disruption after age 12 performed significantly better on the Gi scale than did females regardless of home type.

Two widely held beliefs about the effects of home disruption were not supported by this study. The results of the So scale indicated that male and female children of broken homes are not affected differentially to the extent commonly supposed. In addition, no support was found for the common assertion that boys develop feminine characteristics in a father-absent home environment.

Future research should attempt to control the socioeconomic status of the various home types as well as measure the nature of the present home climate and the type of relationship the child had developed with the remaining parent. The parents' personality as a variable could also be more directly assessed and related to the overall findings.

Summary

The data of this study tended to support the assertion that children from homes broken by separation, divorce, or death are less well-adjusted in terms of the CPI scales of Self-Control, Socialization, Femininity, and Good Impression, than children from intact homes. Girls from broken homes were not found to be more maladjusted than boys from broken homes. No clear cut trend emerged which directly linked the age at which home disruption occurred and the degree of maladjustment.

References

Cox, M. J., & Cox, R. D., Socialization of young children in the divorced family. *Journal of Research and Development in Education*, 1979, *13*, 58-67.

Hammond, J. M. Children of divorce: A study of self-concept, academic achievement, and attitudes. *Elementary School Journal*, 1979, *80*, 55-62.

Hetherington, E. M., Cox, M., & Cox, R. The aftermath of divorce. In J. H. Stevens & M. V. Mathews (Eds.), Mother-child, father-child relations. *National Association for the Education of Young Children*, Washington, D.C., 1978.

Kalter, N. Children of divorce in an outpatient psychiatric population. *American Journal of Orthopsychiatry*, 1977, *47*, 40-51.

Kelly, J. B., & Wallerstein, J. S. The effects of parental divorce: Experiences of the child in early latence. *American Journal of Orthopsychiatry*, 1976, *46*, 20-32.

Rubin, L. D., & Price, J. H. Divorce and its effects on children. *Journal of School Health*, 1979, *49*, 552-556.

Wallerstein, J. S., & Kelly, J. B. The effects of parental divorce: Experience of the preschool child. *Journal of American Academy of Child Psychiatry*, 1975, *14*, 600-616.
Winer, B. J. *Statistical principles in experimental design.* New York: McGraw-Hill, 1962.

The Alberta Journal of Educational Research *Vol. XXVIII, No. 2, June, 1982, 100-112*

EDWARD G. SUMMERS
University of British Columbia

and

J. VIVIAN McCLELLAND
School District #38 (Richmond)

A Field-Based Evaluation of Sustained Silent Reading (SSR) in Intermediate Grades

In this study an eight-stage model was developed for implementing a five-month program of Sustained Silent Reading (SSR) and evaluated using (1) standardized and informal tests, (2) interviews, (3) monitoring/observation, and (4) summative staff ratings. With individual students as the unit of analysis, data from 65 intact treatment and control intermediate grade classrooms (from nine intact elementary schools) revealed no significant differences for covariance adjusted mean scores from standardized and informal reading achievement and attitude measures and no significant interaction effects for reading achievement, attitude, grade level, and sex. However, in the summative evaluation, staff in the SSR schools gave it a strong rating as a desirable language arts activity for incorporation in intermediate grade classrooms.

The starting point of any curriculum evaluation is the question of whether or not a particular curriculum innovation "works." In this research, an eight-stage model for Sustained Silent Reading (SSR) was developed and implemented for five months in intermediate classrooms and evaluated using a variety of measures and data sources. Primary outcome variables centered on the effect of SSR on reading achievement and attitude toward reading.

In 1970 Hunt described a rather unique approach to the stimulation of reading behavior called Uninterrupted Sustained Silent Reading (USSR). In USSR

J. Vivian McClelland is Supervisor of Language Development for School District #38, Richmond, British Columbia.

Dr. Edward G. Summers (PhD, University of Minnesota) is currently Professor of Education in the Department of Language Education, University of British Columbia, where he teaches undergraduate and graduate courses in secondary reading, research, and information retrieval. His special areas of interest include measurement in reading, text structure and discourse analysis, reading and affect, and secondary, college, and adult reading instruction and research.

children are provided with an opportunity to read on their own using books they have selected. The heart of a USSR program requires that children sustain their reading behavior within the structure of a simple set of rules which govern the conduct of a special period of time set aside during the school day for 15 to 25 minutes' duration. Teachers and school staff are all encouraged to participate and provide a model of reading behavior. The reading period is often scheduled school-wide.

Recent adaptations of USSR have utilized various acronyms and proponents such as Allington (1975, 1977), Gambrell (1978), Ganz and Theofield (1974), McCracken (1971), McCracken and McCracken (1972, 1978), Mork (1972), Oliver (1970, 1976), Petre (1971), and Towner and Evans (1975) have described programs operated in both elementary and secondary schools. The notion of sustained silent practice has also been extended to other language arts activities (see Oliver, 1970; Allington, 1975; Cunningham & Cunningham, 1976) and recent writing has even advocated use of the sustained reading practice model by parents in the home.

Programs of USSR have proliferated and many school districts in Canada and the United States have implemented versions of sustained practice. The technique is increasingly described in textbooks and curriculum guides at all levels. Amid this surge of activity, reports often give high marks to the benefits of USSR on reading achievement and attitude toward reading, but such reports tend to be ex cathedra, provide minimal data on restricted populations, and ignore valid research methodology and program evaluation paradigms. Some research has been reported (Cline & Kretke, 1980; Evans & Towner, 1975; Hanson, 1972; Mikulecky & Wolf, 1977; Minton, 1980; Oliver, 1970; Petre, 1971; Sadoski, 1980; Wilmot, 1975; and Wolf & Mikulecky, 1978), but further study is needed to adequately define and develop unbiased measures of the true effect of USSR programs using concepts from current program evaluation methodology. Most of the reported research thus far involves vague definition, strong promises, limited application of evaluation methodology, and weak estimates of effect. The "What would happen if . . ." question posed by our study can be phrased as follows:

> What effect would a carefully conceived and implemented program of Sustained Silent Reading (SSR) have on reading achievement and attitude toward reading scores of a group of intermediate grade students; would the results differ by grade level and sex?

Three null hypotheses related to the research question were established for statistical evaluation.

1. There will be no statistically significant differences on posttest reading and attitude measures (standardized and informal tests) in comparing SSR children with non-SSR children with posttest mean scores adjusted for pretest differences through covariance techniques.
2. There will be no significant interaction effects in comparing Grade 5, 6 and 7 scores by treatment.
3. There will be no significant interaction effects in comparing male and female scores by treatment.

A unique feature of the formative and summative (Scriven, 1974) evaluation project was that it brought together university faculty and public school teachers, administrators, consultative staff, and others in a team research effort aimed at determining the benefits of SSR in a controlled, field-based evaluation project. The full report for the research can be found in Summers (1979).

Evaluation Design

The benefits of random assignment in obtaining an unbiased estimate of program effects have been well documented (Boruch, McSweeny & Soderstrom, 1978; Cook, Fay & Mark, 1977). However, true random assignment is seldom possible in most field-based research settings necessitating less controlled quasi-experimental and ethnographic approaches to evaluation. Thus, our research was field-based and quasi-experimental consisting of treated and nontreated groups, nonrandom assignment of groups, and pretest and posttest measures. We made every effort to maximize the internal validity of the study. The research best fits Campbell and Stanley's (1966) ubiquitous nonequivalent control group design which is thoroughly discussed by Cooke and Campbell (1976), Fitz-Gibbon and Morris (1978), Patton (1980), and Rossi and Wright (1977) in terms of models for program evaluation in field settings. Analysis of variance and covariance were used to assess differences between the treated (SSR) and untreated (non-SSR) groups (Kenny, 1975; Wildt & Ahtola, 1978). Type 1 error was set at .05.

Schools

Nine schools located in suburban Richmond, British Columbia were used in the study. Four elementary schools, judged to be representative of those in the district, were selected as SSR schools. Each of the schools contained approximately 10 classrooms at the grade 5, 6, and 7 levels. A total of five schools were selected from those remaining to serve as the non-SSR comparison group and provide the requisite control classrooms at the grade 5, 6, and 7 levels. The study ultimately involved close to 1,400 children, approximately 25% of the total intermediate group in the district, in nine intact schools and 65 intact classrooms. The SSR and non-SSR groups were carefully "balanced" by district staff to ensure good equivalency on variables such as teaching staff, student ability, socioeconomic level, type of instructional program, reading achievement and general achievement, attitude toward school, library facilities, and home environment. In comparing the entry level reading and attitude scores for the SSR and non-SSR groups, data from 60 single factor analyses of variance, using seven reading achievement and reading attitude pretest mean scores in treatment, grade, and gender comparisons revealed only one significant F statistic ($p<.05$), with two approaching significance on the initial scores. It was concluded that good equivalency was achieved in forming the treatment and comparison groups. It was also revealed that scores fell across the entire range of reading achievement and attitude measures closely approximating a normal distribution for both groups. Data on teacher qualification level, sex, years of experience, and academic degrees provided an indication that the 75 teachers and principals in the SSR and non-SSR groups were very comparable on these variables.

Non-SSR schools followed their regular programs but received exactly the same testing as the SSR schools. Thus, the regular programs taking place in the non-SSR schools became the competing model to the SSR program. SSR was implemented on an all-school basis in all grades in the treatment schools but the evaluation took place only in grades 5, 6, and 7. The treatment was implemented for a period of approximately five months, December through June 1978.

Statistical Analysis

A major question in all program evaluation research concerns the specification of an appropriate unit of analysis. Several levels of data aggregation are possible including the individual, the class, the school, the district, or beyond.

Individuals were considered the unit of analysis in this study. We argued that the SSR program was being given to children and that children were responding independently of other children vis à vis the individualized nature of the SSR activity. It must be admitted, however, that children work within classroom and school environments which can influence achievement and attitude; aggregation by these large units could also be justified in determining differential treatment effects. However, the classroom and school effects were felt to be "balanced" by the careful selection of the comparison SSR and non-SSR groups. An excellent discussion of the overall issue of school and classroom (teacher) effect was recently reported by Centra and Potter (1980).

Various statistical packages were used to generate descriptive data (SPSS, BMD). The *Multivariance Version VI* program by Finn (1977) was used to provide contrasts to test the null hypotheses. The model for the study resembles a typical A (treatment) x B (grade) x C (sex) factorial design. However, it is different in that many of the comparisons that would be available in the complete design were not examined. We felt there was no point in including results from a three-factor interaction if we, first of all, would not initially hypothesize such an effect, and secondly, if its interpretation would not add to our understanding of the results. The multivariance program allows for manipulation of nested data in this fashion. Thus, multivariance analysis of single effects, and their interaction with factors of particular interest (using covariates to increase the power of the statistical analysis), seemed the most appropriate approach (see Huck [1972] and Games [1978] for discussion of the points which influenced the development of the statistical analysis model).

Measuring Instruments

In general, good program evaluation should involve multiple data sources (Anderson & Ball, 1978) and systematic examination of rival hypotheses in interpreting the results (Scriven, 1976). Well defined evaluation research also requires careful conceptual and operational specification of the major relevant variables.

The outcome domains that seemed most desirable and practical to assess were those which could be conveniently measured in a group testing situation using recognized batteries of standardized tests appropriate for preadolescent children. The state-of-the-art to date (as represented by the literature reviewed), in assessing the effects of SSR programs, suggests that an analysis of the effect on reading achievement and attitude toward reading would be most appropriate.

Reading comprehension was defined as "obtaining meaning from words and larger units of discourse in printed graphic stimuli" and was assessed with the vocabulary and comprehension sections of the Metropolitan Achievement Test, Intermediate and Advanced forms (Durost, Bixler, Wrightstone, Prescott & Balow, 1970-71). The MAT is a widely used achievement battery. The publisher's reliability estimates were accepted for the test. Attitude toward reading was defined as "the tendency to respond in a favourable or unfavourable manner to specific questions about reading" and was measured using the 15 reading items of the Estes Attitude Scale (Estes, Johnstone & Richards, 1975). The selection of the Estes was based on a careful review of existing attitude scales (Summers, 1977b). Hoyt analysis of variance measures (Nelson, 1974) indicated good internal consistency for the Estes Reading Attitude Scale by grade level and gender within grade (grade 5, $r=.86$; grade 6, $r=.87$; grade 7, $r=.89$). Data approximated a normal distribution and total and subscale correlations were in the expected

direction. Complete validation data for the Estes Reading Attitude Scale for intermediate grades, using a Canadian sample, is reported in Summers (1980).

An informal measure of reading achievement and attitude was also developed to provide cross-validation results for the standardized achievement and attitude measures. The Student Reading Rating Scale (SRRS) provides a measure of the child's subjective perception of his reading achievement, liking for reading, and the quantity of reading done. Children are asked to rate themselves, in comparison with peers, using a five-point Likert scale on three questions: (1) About how *Well* do you read? (2) How well do you *Like* to read? (3) About how *Much* do you read? The range of scores is 3 to 15 with an approximate midpoint of 9. Hoyt analysis of variance measures also indicated good internal consistency for the SRRS informal measure by grade level and gender within grade (grade 5, $r=.76$; grade 6, $r=.84$; grade 7, $r=.89$). Data approximated a normal distribution and total and subscale correlations were in the expected direction.

Definition and Implementation of SSR

In evaluation research, the basic assumption is that the educational treatment itself, its goals, and the criteria for implementation are sufficiently well defined so as to allow an appropriate evaluation to be designed and carried out (Leinhardt, 1980).

If the program to be implemented is not carefully defined "elaborately designed evaluation studies may sometimes end up appraising non-events with no one the wiser" (Charters & Jones, 1973, p. 5). Ideas from the Fullan and Pomfret (1977) review guided development of the plan for program implementation of SSR. Techniques described in Fink and Kosecoff (1980), Fitz-Gibbon and Morris (1978), House (1980), Leithwood and Montgomery (1980), and Morris and Fitz-Gibbon (1978) provide good information for program design, measurement of innovations, and evaluation methodology.

Hall and Loucks (1977) conceptualized program implementation in terms of "levels of use" or stages of growth toward some ideal of the innovation being implemented. Eight dimensions comprise their levels of use paradigm: (1) Nonuse, (2) Orientation, (3) Preparation, (4) Mechanical Use, (5) Routine, (6) Refinement, (7) Integration, and (8) Renewal. The levels have research validation and are described more fully in Loucks, Newlove and Hall (1975) and Hall, Loucks & Rutherford (1975). We felt the generic concept of levels of use was a good technique for visualizing program definition and, most importantly, the eight stages provided a means of incrementing the SSR program implementation for later analysis and evaluation. Figure 1 presents our adaptation of the eight Hall and Loucks use levels in defining the stages of implementation for the SSR project.

The complete description of the SSR program which was implemented can be found in Summers (1977a, 1979) and Summers and Snyder (1978). The basic concepts were: (1) schoolwide scheduling, (2) modelling of reading by school staff, (3) all children reading self-selected material, (4) ready access to reading material, (5) protection of SSR periods from outside interruption, (6) consistent praise and positive feedback for reading efforts, (7) avoidance of obtrusive record keeping, (8) careful scheduling of SSR periods and consistent time limits, and (9) incremental adoption beginning with small periods spaced throughout the week and moving to 20- to 25-minute periods four or five times a week.

Two released time workshops were scheduled for school staffs to begin the project. *Orientation Workshops* discussed the goals, reviewed pertinent literature,

presented definitions, described use of instruments and set up scheduling and testing. *Implementation Workshops* reviewed goals, discussed the relationship between SSR, reading interest, and the school library, worked out implementation details, defined SSR periods for each school, and established monitoring, interviewing and posttesting schedules. Pretests were administered in December and January and posttests were administered in June 1978. All testing was done by principals, vice-principals, and teachers.

Stage of use	Definition of activities	Decisions
Nonuse	Stage at which school staff have minimal knowledge of SSR as a concept, little involvement with the procedure on a systematic basis, and are not directly moving to adoption of SSR on a total school basis.	Decision Point A Take action to learn more about SSR and explore its adoption for classroom use.
Orientation	Stage at which school staff acquire information about SSR, explore its value as a reading practice and motivation technique, and determine the feasibility of a tryout period in the classroom.	Decision Point B Decision is made by school staff to implement SSR on a trial basis.
Preparation	Stage at which school staff prepare for use of SSR, plan instructional procedures and details of implementation. Plan for coordination between teachers, librarians, and administrators.	Decision Point C School district provides the resources, inservice time, and materials necessary to insure that school staff have requisite understanding and skills to implement SSR on a trial basis.
Mechanical Implementation	Stage at which school staff focus effort on gradual introduction of SSR, mechanics of implementation, short term day-by-day use and adjustment to classroom, school, and library needs. Focus is on step-by-step mastery of tasks, by children and teacher, for effective introduction and implementation of SSR. Coordination between teachers, library and administrators is complete. Adult reading modelling is initiated on a day-by-day basis.	Decision Point D A routine pattern of use by school staff has been established for SSR.
Routine Implementation	School staff use of SSR is established, scheduling pattern is stabilized, few if any changes are being made in ongoing use. Adult reading modelling is refined to a consistent activity.	Decision Point E School staff reflect on use of SSR and its consequences, informal evaluation to develop more efficient implementation and improve student benefits.
Refinement	Stage at which school staff supplement and extend use of SSR throughout language arts activities within classroom and school environments based on long and short term consequences for students. Adult reading modelling continues.	Decision Point F School staff initiate refinements and modification in use of SSR and coordinate with colleagues for maximum benefit.
Integration	Stage at which school staff are combining own efforts with colleagues, using related activities, to enhance collective impact on students. Adult reading modelling continues.	Decision Point G Continued refinement and modification in coordination and use for maximum effect.
Adoption	Stage at which decision is made to implement SSR as a regular language arts activity on a scheduled school-wide basis.	Decision Point H SSR becomes a regular activity in elementary schools on the recommendation of the school staff and principals.

Figure 1. Incremental stages in SSR implementation and adoption.

Focused interviews (Foster & Nixon, 1975; Loucks, Newlove & Hall, 1975; Weiss, 1975) were held in March and April, 1978, with the participating teachers, librarians, and principals in all the SSR schools. The 20-minute interviews were

individual and relaxed but did focus on obtaining specific answers to a series of questions related to: (1) feedback on the mechanics of implementing SSR, (2) reaction of children to SSR, (3) library utilization, (4) home response, (5) aspects of all school scheduling, (6) analysis of pretesting, (7) reaction to the orientation and implementation workshops, (8) expressed needs and problems in relation to the project, and (9) previewing of upcoming posttesting activities. The summative questionnaire was filled out by participating SSR staff at the conclusion of the program in June, 1978.

Results

The statistical design allowed us to test for the equality of posttest means after extracting the variance which could be explained by pretest scores.

The contrasts for the first null hypothesis required comparing posttest mean raw scores for the two treatment groups (SSR, $n=612$; non-SSR, $n=630$) with raw scores aggregated across gender, class, grade, and school for both groups, with pretest scores used as covariates in all comparisons. *No significant F statistics ($p<.05$) were found for any of the six standardized and informal reading achievement and reading attitude scores and we failed to reject the null hypothesis for all six dependent variables.* (All contrasts have 1, 1106 df; MAT Vocabulary $F=.49$, n.s., $p<.52$; MAT Comprehension $F=1.82$, n.s., $p<.17$; Estes $F=.62$, n.s., $p<.16$; SRRS *Well* $F=.74$, n.s., $p<.38$; SRRS *Like* $F=2.89$, n.s., $p<.08$; SRRS *Much* $F=.94$, n.s., $p<.22$.)

The second null hypothesis contrasted SSR ($n=305$) with non-SSR ($n=306$) for males, relative to SSR ($n=306$) with non-SSR ($n=325$) for females on posttest mean raw scores, with pretest covariance adjustments, to determine if one treatment was more effective than the other for gender. *No significant F statistics were found and we failed to reject the treatment by gender interaction null hypothesis for all six dependent variables.* (All contrasts have 1, 1106 df; MAT Vocabulary $F=.54$, n.s., $p<.46$; MAT Comprehension $F=2.73$, n.s., $p<.09$; Estes $F=.37$, n.s., $p<.54$; SRRS *Well* $F=1.02$, n.s., $p<.31$; SRRS *Like* $F=.00$, n.s., $p<.96$; SRRS *Much* $F=.34$, n.s., $p<.55$.)

The third null hypothesis required assessing the interaction of treatment with grade level, with pretest scores used as covariates in all comparisons, to determine if one treatment was more effective than the other grade by grade. *There was one significant F statistic for the interaction between treatment effects and the contrast of grade 5 with grades 6 and 7 combined* (all contrasts have 1, 1106 df, MAT Vocabulary $F=.19$, n.s., $p<.65$; MAT Comprehension $F=12.68$, $p<.0004$; Estes $F=1.25$, n.s., $p<.26$; SRRS *Well* $F=1.19$, n.s., $p<.27$; SRRS *Like* $F=.70$, n.s., $p<.40$; SRRS *Much* $F=1.94$, n.s., $p<.16$) *and no significant F statistics for the interaction between treatment effects and the contrast of grade 6 with grade 7* (MAT Vocabulary $F=1.09$, n.s., $p<.24$; MAT Comprehension $F=1.60$, n.s., $p<.10$; Estes $F=1.88$, n.s., $p<.16$; SRRS *Well* $F=1.12$, n.s., $p<.28$; SRRS *Like* $F=.77$, n.s., $p<.38$; SRRS *Much* $F=.04$, n.s., $p<.82$). *The one significant F for MAT Comprehension is inexplicable. Other than that one contrast, we failed to reject the treatment by grade level null hypothesis for all six dependent variables.* Tables of covariance adjusted mean raw scores and full data for all statistical tests are available from the authors.

Discussion

The null is always a difficult hypothesis to reject and this study proved no exception. The statistical results are unequivocal: *our data revealed no significant*

differences in average covariance adjusted mean scores for the SSR treatment and non-SSR comparison groups using both standardized and informal measures of reading achievement and attitude toward reading. The results for the treatment by grade and treatment by gender interactions were also nonsignificant. In field research such as this, the interpretations of the statistical results must include consideration of possible alternative explanations.

1. It could be argued that the obtained results are inaccurate because the treatment and comparison groups were not initially equivalent in reading achievement and attitude toward reading. We would reject this as an alternative explanation.
2. Poor implementation of the SSR treatment could have affected outcomes. Any curriculum evaluation activity is, ultimately, a test of the school district's ability to conceptualize and conduct a research project. Because of the careful monitoring involved, we would reject poor treatment implementation as an alternative explanation.
3. One alternative explanation that has some credibility is that the SSR treatment was not of sufficient duration to have maximum effect. A five-month trial of an all school effort is, however, a substantial treatment period in comparison with other SSR evaluations to date.

   Some evidence is available which supports the notion that a longer treatment period might have been desirable if we examine the incremental levels of use outlined in Figure 1. Different schools could have achieved different levels of use. Teachers could also vary in the extent to which they implemented SSR. The focused interviews and school visits provided data on the start-up and implementation of the project. Start-up activities, in particular, take longer than anticipated, especially when total school efforts are involved. It is estimated that, overall, the SSR implementation moved slightly beyond the Routine Implementation stage during the five-month trial.
4. It could be argued, with some validity, that weak adult modelling of reading influenced the level of use attained in the study. The focused interviews revealed that teachers were somewhat reluctant to read during SSR periods, especially in the initial phases of the project. It is subjectively estimated that only 65 to 70 percent of the total modelling that could have been done by adults during the course of the study actually took place. It has been recently suggested (McCracken & McCracken, 1978), that adult modelling may be "the" crucial element in the success of SSR. Thus, socialization of teachers into what may be for them a new role becomes an important element in implementing SSR programs. The implementation of our SSR project, of necessity, followed more the fidelity than the mutual adaptation model as outlined by Fullan and Pomfret (1977). The mutual adaptation model would provide a better framework for socialization and role change and should be considered for projects like SSR. While it is impossible to determine the net statistical effect, it is very likely that limited modelling could have influenced the outcome results.
5. The selection of criterion measures is critical in the evaluation of curriculum innovation. The measures should be psychometrically sound and reflect the objectives of the project. Although different batteries of tests could have been selected, those used in the project are well accepted as measures of reading achievement and attitude. The SRRS also served to cross-validate the results from the standardized tests. In all instances, the instruments used either

have a history of good reliability or such reliability was supported by the data gathered in the study.

6. Although the measurement/objectives mismatch hypothesis in (5) above is untenable, we have to entertain the possibility of Type II error with respect to the measurement of the SSR effect. In essence, we may have failed to reject the null hypotheses, based on data from the instruments used to measure the SSR effect, when in fact we might have rejected them had other instruments been used. We are sensitive to the fact that achievement and attitude can be measured differently and that a range of outcomes beyond achievement and attitude could also be measured in evaluating SSR. However, we also cannot avoid the fact that little consideration has thus far been given in the literature to development of an adequate conceptual framework which would allow us to predict the true effects of SSR programs. As yet, we have only obscure knowledge of the full characteristics of the subtleties of SSR as an independent variable. This makes it difficult to select or develop criterion measures beyond rather gross available standardized and informal tests. The prevailing view is to assume that sustained practice will have an effect on achievement and attitude. However, it is possible that effects do exist which were not being measured by the tests used.

   In attempting to develop a conceptual framework to account for the possible effects of SSR, it is highly likely that the majority of benefits attributable to such programs may never be amenable to statistical evaluation. For example, it is often suggested that scheduling SSR periods on a schoolwide basis emphasizes that the school considers practice of reading and reading for pleasure as being just as important as reading pursued as an academic activity. Placing emphasis on these aspects of reading is a legitimate school objective and their attainment can have far reaching consequences.

   At present, districts which feel they wish to incorporate SSR programs should proceed to do so because of the inherent value of SSR itself rather than the expectation that SSR will have strong measured effects on achievement and attitude when compared to competing program X or Y. The catalytic effect which SSR can have on children and teachers, in terms of developing interest and providing in-school practice of reading, may ultimately be of far greater importance than any precisely measured statistical outcomes.

7. One final alternative explanation should be considered. Our model is not the only possible design for evaluating SSR. Program evaluation design is both a science and an art and design/analysis mismatch is a possibility in any study.

   Five problems loom large in any curriculum evaluation: (1) selection of an appropriate unit of analysis, (2) the intact, hierarchical organization of school populations requiring rather complicated designs when evaluating across units, (3) use of covariance techniques in quasi-experimental field setting, (4) nonrandom formation of treatment groups, and (5) selection or development of adequate criterion measures. Replication using another design or reanalysis of our data might provide interesting and different results. We would welcome such activity.

In considering the above seven alternative explanations, we would conclude that our treatment groups were sufficiently equivalent, the SSR treatment was reasonably well implemented, measuring instruments were appropriate to the

objectives of the study and psychometrically sound, and the design and statistical analysis were reasonable for field research activities. In contrast, a more extensive treatment period might be desirable, adult modelling could have been stronger, role change and socialization processes should perhaps have received greater emphasis, and criterion measures other than those used might have produced different results. In summary, although the alternative explanations all have validity, we conclude that we obtained a reasonably valid and unambiguous test of the null hypotheses.

Conclusion

The state-of-the-art in curriculum evaluation includes dozens of models which can be applied in program evaluation. We followed a multimeasure (Tuckman, 1975), somewhat eclectic approach: we derived statistical hypotheses and evaluated outcome scores using both standardized and informal measures; we evaluated through observation-monitoring and personal focused interviews; and we used participant summative rating scales, all in an attempt to gain insight into the effect of SSR programs. The reader can judge the efficacy of our approach.

We found the Fullan and Pomfret conceptualization of factors to be considered in implementing curriculum change to be particularly useful in the SSR evaluation. Their analysis makes interesting reading for any school district involved in evaluating reading programs. The Hall and Loucks model provided a good paradigm for incrementing stages in curriculum implementation and we also recommend it for use in curriculum evaluation. The Morris and Fitz-Gibbon (1978), Fitz-Gibbon and Morris (1978), Patton (1980), and House (1980) texts all provide good examples of useful information in structuring and carrying out a program evaluation. The concept of the focused interview has considerable merit. In relatively short interviews (15 to 20 minutes), using preset questions, we were able to gather much valuable data from participating teachers, principals, and librarians in the SSR program.

As is the case in much curriculum innovation, the teachers implementing the program provide perhaps the most important input in estimating the worth of an activity. Thus, although the *statistical* results for the study were nil, interesting *educational* implications exist. Our final summative rating questionnaire revealed that the 37 teachers, librarians, and principals in the SSR schools reacted very favorably to SSR as a curriculum activity in intermediate grades. (1) They strongly favored all school scheduling — four to five times a week, SSR periods beginning either the morning or afternoon school session, and room libraries in support of SSR activities. (2) They rated children's reactions as being very positive overall and suggested that SSR be implemented in junior secondary grades as well. (3) Two-thirds of the respondents rated SSR as being very beneficial in developing reading interests, and almost all rated some increase in the range of topics read as a result of SSR in their classrooms. (4) The respondents, almost without exception, subjectively rated SSR as influencing development of positive attitude toward reading as a school subject, suggested it would benefit the total range of children in intermediate classrooms, and recommended that SSR be adopted as a regular language arts activity in the elementary schools of the district. These subjective data may provide some of the strongest support for adoption of SSR programs.

Appreciation is expressed to the Faculty of Education, University of British Columbia for computer and statistical services in completing this project and to the Faculty of Education, University of Western Ontario for secretarial services to complete the report during the sabbatical leave of the senior author.

References

Anderson, S., & Ball, S. *Profession and practice of program evaluation.* San Francisco: Jossey Bass, 1978.

Allington, R. Sustained approaches to reading and writing. *Language Development*, 1975, *52*, 813-815.

Allington, R. If they don't read much, how they ever gonna get good? *Journal of Reading*, 1977, *21*, 57-61.

Boruch, R. A., McSweeny, A., & Soderstrom, E. Randomized field experiments for program planning, development and evaluation. *Evaluation Quarterly*, 1978, *4*, 655-695.

Campbell, D. T., & Stanley, J. C. *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally, 1966.

Centra, J. A., & Potter, D. A. School and teacher effects: An interrelational model. *Review of Educational Research*, 1980, *50*, 273-291.

Charters, W. W., Jr., & Jones, J. E. On the risk of appraising non-events in program evaluation. *Educational Researcher*, 1973, *2*, 5-7.

Cline, R. K. J., & Kretke, G. L. An evaluation of long-term SSR in the junior high school. *Journal of Reading*, 1980, *23*, 503-506.

Cook, T. D., Fay, L. C., & Mark, M. M. Randomized and quasi-experimental designs in evaluation research: An introduction. In L. Rutman (Ed.), *Evaluation research methods: A basic guide.* Beverly Hills: Sage Publications, 1977.

Cooke, T. D., & Campbell, D. T. The design and conduct of quasi-experimental and true experiments in field settings. In M. D. Dunnette (Ed.), *Handbook of industrial and organizational psychology.* Chicago: Rand McNally, 1976.

Cunningham, P. M., & Cunningham, J. W. SSSW better content writing. *Clearinghouse*, 1976, *49*, 237-240.

Durost, W. N., Bixler, H. H., Wrightstone, J. W., Prescott, G. A., & Balow, I. H. *Metropolitan achievement tests: Teacher's handbook, teacher's directions, intermediate and advanced forms.* California: Harcourt Brace Jovanovich Inc., 1970-1971.

Estes, T. E., Johnstone, J. P., & Richards, H. C. *Estes attitude scales.* Charlottesville, Virginia: Virginia Research Associates Ltd., 1975.

Evans, H. M., & Towner, J. C. Sustained silent reading: Does it increase skills? *Reading Teacher*, 1975, *29*, 155-156.

Fink, A., & Kosecoff, J. *An evaluation primer.* Beverly Hills: Sage Publications, 1980.

Finn, J. D. *Multivariance Version VI.* Chicago: International Education Services, 1977.

Fitz-Gibbon, C. T., & Morris, L. L. *How to design a program evaluation.* Beverly Hills: Sage Publishers, 1978.

Foster, L., & Nixon, M. The interview reassessed. *The Alberta Journal of Educational Research*, 1975, *21*, 18-22.

Fullan, M., & Pomfret, A. Research on curriculum and instruction implementation. *Review of Educational Research*, 1977, *47*, 335-397.

Gambrell, L. B. Getting started with sustained silent reading and keeping it going. *Reading Teacher*, 1978, *32*, 328-331.

Games, P. A. Nesting, crossing, Type IV errors, and the role of statistical models. *American Educational Research Journal*, 1978, *15*, 253-258.

Ganz, P., & Theofield, M. B. Suggestions for starting SSR. *Journal of Reading*, 1974, *17*, 614-616.

Guttentag, M., & Struening, E. L. (Eds.), *Handbook of educational evaluation research, Vol. 1 and 2.* Beverly Hills: Sage Publications Inc., 1975.

Hall, G. E., Loucks, S. F., & Rutherford, B. W. Levels of use of the innovation: A framework for analyzing innovation adoptions. *Journal of Teacher Education*, 1975, *26*, 52-56.

Hall, G. E. The study of individual teacher and professor innovations. *Journal of Teacher Education*, 1976, *27*, 22-23.

Hall, G. E., & Loucks, S. F. A developmental model for determining whether the treatment is actually implemented. *American Educational Research Journal*, 1977, *14*, 263-276.

Hanson, R. A. A study of the effects of SSR on the reading attitude and habits of second, third and fourth grade students in a low income area. (Unpublished doctoral dissertation, University of North Dakota, 1972.) *Dissertation Abstracts International*, *33*, 5034A.

House, E. *Evaluating with validity*. Beverly Hills: Sage Publications, 1980.

Huck, S. W. The analysis of covariance: Increased power through reduced variability. *The Journal of Experimental Education*, 1972, *41*, 42-46.

Hunt, L. The effects of self-selection, interest and motivation upon independent, instructional and frustrational levels. *Reading Teacher*, 1970, *24*, 146-151.

Kenny, D. A. A quasi-experimental approach to assessing treatment effects in the nonequivalent control group design. *Psychological Bulletin*, 1975, *82*, 345-362.

Leinhardt, G. Modeling and measuring educational treatment in evaluation. *Review of Educational Research*, 1980, *50*, 393-420.

Leithwood, K. A., & Montgomery, D. J. Evaluating program implementation. *Evaluation Review*, 1980, *4*, 193-214.

Loucks, S. F., Newlove, B. W., & Hall, G. E. *Measuring levels of use of the innovation: A manual for trainers, interviewers and raters*. University of Texas, Austin, Texas, 1975.

Mikulecky, L., & Wolf, A. The effect of uninterrupted sustained silent reading and of reading games on changes in secondary students reading attitudes. In P. D. Pearson (Ed.), *Reading: Theory, research and practice*. Twenty-sixth Yearbook of the National Reading Conference. Clemson, South Carolina: National Reading Conference, 1977.

Minton, M. J. The effect of Sustained Silent Reading upon comprehension and attitudes among ninth graders. *Journal of Reading*, 1980, *23*, 498-502.

Mork, T. SSR in the classroom. *Reading Teacher*, 1972, *25*, 438-441.

Morris, L. L., & Fitz-Gibbon, C. T. *How to measure program implementation*. Beverly Hills: Sage Publications, 1978.

McCracken, R. A. Initiating sustained silent reading. *Journal of Reading*, 1971, *14*, 521-524, 582-583.

McCracken, R. A., & McCracken, M. J. *Reading is only the tiger's tail*. San Rafael, California: Leswing Press, 1972.

McCracken, R. A., & McCracken, M. J. Modelling is the key to sustained silent reading. *Reading Teacher*, 1978, *30*, 406-408.

Nelson, L. R. *Guide to LERTAP use and interpretation*. Dunedin, New Zealand: University of Otago, 1974.

Oliver, M. E. High intensity practice; the right to enjoy reading. *Education*, 1970, *91*, 69-71.

Oliver, M. E. Effect of high intensity practice on reading achievement. *Reading Improvement*, 1976, *13*, 226-228.

Patton, M. Q. *Qualitative evaluation methods*. Beverly Hills: Sage Publications Inc., 1980.

Petre, R. M. Reading breaks make it in Maryland. *Journal of Reading*, 1971, *15*, 191-194.

Rossi, H., & Wright, S. R. Evaluation research: An assessment of theory, practice, and politics. *Evaluation Quarterly*, 1977, *7*, 5-52.

Sadoski, M. C. An attitude survey for sustained silent reading programs. *Journal of Reading*, 1980, *23*, 721-726.

Scriven, M. Evaluation perspectives and procedures. In W. Popham (Ed.), *Evaluation in education*. Berkeley: McCutchan, 1974.

Scriven, M. The methodology of evaluation. In R. W. Tyler et al. (Eds.), *Perspectives of curriculum evaluation*. Skokie, Illinois: Rand McNally, 1976.

Summers, E. G. *Implementing Sustained Silent Reading in typical teaching situations*. Vancouver: University of British Columbia, 1977. (a)

Summers, E. G. Instruments for assessing reading attitudes: A review of research and bibliography. *Journal of Reading Behavior*, 1977, 9, 137-165. (b)

Summers, E. G. An evaluation of the effects of a program of *Sustained Silent Reading* (SSR) on reading achievement and attitude toward reading in intermediate grades (final report), 1979. (ERIC Document Reproduction Service ED 169 473.)

Summers, E.G. The validity of the *Estes Reading Attitude Scale* for intermediate grades. *The Alberta Journal of Educational Research*, 1980, *26*, 36-43.

Summers, E. G., & Snyder, G. *SSR, reading interests and the school library*. Vancouver: University of British Columbia, 1978.

Towner, J. C., & Evans, H. M. The sustained silent reading: Does it float? *Reading Horizons*, 1975, *15*, 83-87.

Tuckman, B. W. The differential outcome hypothesis. *CEDR Quarterly*, 1975, *11*, 16-18, 21.

Weiss, C. H. *Interviewing in evaluation research. Handbook of evaluation research, Vol. 1.* M. Guttentag & E. L. Struening, Eds. Beverly Hills: Sage Publications, 1975.

Wildt, A. R., & Ahtola, O. T. *Analysis of covariance*. Beverly Hills: Sage Publications, 1978.

Wilmot, M. P. An investigation of the effect upon reading performance and attitude toward reading of elementary grade students of including in the reading program a period of sustained silent reading. 1975 (ERIC Document Reproduction Service No. ED 119 130).

Wolf, A., & Mikulecky, L. Effects of uninterrupted sustained silent reading and reading skills instruction on changes in secondary students' reading attitudes and achievement. Paper presented at the National Reading Conference, New Orleans, 1978.

The Alberta Journal of Educational Research *Vol. XXVIII, No. 2, June, 1982, 113-121*

RODNEY V. McCORMICK
Southern Illinois University at Carbondale

The Balance Effect on the M-Space by Spatial Interaction

A theoretical foundation is given for the concept of Balance which is related to the Piagetian concept of cognitive equilibrium. A discussion is given concerning possible interactions of Balance with Spatial Ability and M-space, i.e., the span of short-term memory. Graduate statistics students were subjects in the investigation of these possible interactions. The investigation revealed a sizable Balance by Spatial by M-space triple interaction. A statistical interpretation was given to the results along with a discussion of the concurrence of the results to theory.

The purpose of this study is to make operational the concept of "balance" (Piaget, 1977) and to investigate its interactive effects with spatial ability (McGee, 1979; Smith, 1964) and M-space (Case, 1972, 1974, 1975, 1977; Pascual-Leone, 1970) on graduate level statistics course work.

Piaget's (1977) theory of equilibration describes the dynamic balance of an organism with its environment and that maintained among the subsystems of the organism and its relation to cognitive development. Two postulates of his theory of equilibration are as follows:

"First postulate: Any scheme of assimilation tends to feed itself; that is, to incorporate outside elements compatible with its nature into itself" (p. 7). This suggests that a person by virtue of being alive tends to integrate into his/her action or conceptual schemes those elements that are consistent with the nature of that scheme.

"Second postulate: The entire scheme of assimilation must alter as it accommodates to the elements it assimilates; that is, it modifies itself in relation to the particularities of events but does not lose its continuity. . . nor its earlier powers of assimilation" (p.7). This is essentially a postulate that assimilation and accommodation remain in balance "in order for the accommodation to succeed" (p.8).

Rodney V. McCormick is presently an Assistant Professor of Physics and Educational Research at South Carolina State College. He received a BSc Physics from the University of Illinois, and MSc Physics and PhD Education from Southern Illinois University at Carbondale, where he taught Statistics as Visiting Assistant Professor before joining SCSC. Present research interests include complex interactions in learning.

Piaget suggests that there are three kinds of equilibration, or balancing. The first occurs during a subject-object interaction. "There is the equilibration between the assimilation of schemes of action and the accommodation of these to the objects" (p. 9). The balance then has to be maintained between bringing an object into an action scheme and the modification of the action scheme to the object. Nonbalance occurs in this sense when an object cannot be assimilated into an action scheme.

The second form of balancing occurs during the interactions of the subsystems of the subject; that is, the interaction of action or conceptual schemes. Two or more conceptual or action schemes equilibrate, or come into balance, through reciprocal assimilation and accommodation. Nonbalance is caused here by the fact that "the subsystems are generally constructed at different speeds " (p. 9). One may understand one concept better than a second, and, therefore, missing components of the second make an emerging concept incomplete.

The third form of balancing is that between integration and differentiation of schemes.

> On the one hand, to differentiate a totality, T, in subsystems, S, means not only to confirm what each of these possesses but also to exclude, or deny, the characteristics each subsystem does not have. On the other hand, to form (to integrate) a total system, T, means to free positively the characteristics common to all the S, but this also means to distinguish — this time negatively — the common features of special characteristics not belonging to T. (p. 11)

Both Phillips (1975) and Ginsberg and Opper (1979) have put Piaget's theoretical abstractions into more concrete, lexical, and perhaps measurable terms. Phillips describes the development of play as a process of differentiation of means separate from ends. Phillips states:

> The separation of means from ends has far reaching implications. When the two are completely separated and the ends drops out (the means becomes an end in itself), we have *play*; when they are differentiated but continually related, we have *problem-solving* behavior. Both originate in this primitive separation. (p. 36)

According to Piaget (1970), play is the "assimilation of reality into the self" (p. 22). Ginsberg and Opper (1979) suggest imitation as a concrete example of accommodation. Children may imitate the shape of a fish hook by bending one finger (they have accommodated the finger-bending action scheme to a fish hook). On the other hand, children in play may pretend that two sticks crossed together is an airplane, in which case they have assimilated the object into an airplane scheme. In adults, intellectual play may take the form of "brainstorming" where unusual associations, and diverse ideas are generated, usually by noting similarities between unrelated things. In adults, imitation may take the form of "cookbook" problem solving, where one's actions accommodate directly to step-by-step instructions. Rote memorization becomes simply an imitation of previously perceived material.

It has been attempted to measure via a self-report questionnaire the degree to which assimilation and accommodation are generally in balance among students during the course of studying graduate level statistics. Questions were inspired by Samples' (1968, 1976) balance of metaphorical mode (or playful) vs. substantive mode (or work) in natural problem-solving behavior of children as well as by Piagetian concepts described above. Both Piaget's and Samples' play-like components of problem-solving involve assimilating similarities between the external material and an internal scheme or metaphor. Piaget (1977) suggests that

a nonbalance occurs when "an obstacle" holds an "assimilation in check" (p. 187). One cannot assimilate the elements of the external reality unless one has the systems to do so. Just as the grasping action scheme accommodates to a large object by having the hand open wide, students accommodate their conceptual schemes to extremely difficult subject matter through imitation. They may imitate the problem-solving procedures given by the teacher, take prolific notes as a verbal imitation of the teacher's knowledge, or read passages out of the text numerous times until they are capable of imitating the text verbatim. These methods in lexical terms are simply rote memorization. Questions pertaining to a general nonbalance in favor of accommodation and work stressed this memorization aspect.

Questionnaire

Instructions: Circle the number next to the statement according to the degree of your agreement whether the statement describes you or your feelings.
0= do not agree
1= agree
2= strongly agree

1. I do statistics problems the same way as following a recipe to cook something.	*0*	*1*	*2*
2. When I listen to a lecture in statistics I take a lot of notes so that I can learn the material later.	*0*	*1*	*2*
3. I look forward to statistics class.	*0*	*1*	*2*
4. Statistical ideas are so similar to some non-mathematical things I already know, I feel as if I understand statistics because of this similarity.	*0*	*1*	*2*
5. In order to understand statistics I have to do a lot of memorizing.	*0*	*1*	*2*
6. I am taking statistics more for the reason that I enjoy the material than to fulfill a requirement of my program.	*0*	*1*	*2*
7. Studying statistics is more like playing with ideas than working at a job.	*0*	*1*	*2*
8. I don't take notes in statistics lecture because everything seems to "just fit" into what I already understand.	*0*	*1*	*2*
9. It seems that I have to repeat an idea in statistics many times before I can even begin to understand it.	*0*	*1*	*2*
10. Doing statistics is real work.	*0*	*1*	*2*

A student who takes a very playful approach to the course material, like a child at play, will see similarities in just about everything. But again Piaget (1977)

suggests that the initial cause of the nonbalance is the "systematic primacy of the positive characteristics" and the "assimilation at the beginning of affirmations almost exclusively" (p. 189). That is, a student will be unbalanced by assimilating only the similarities between the course material and the chosen scheme of assimilation. It is on this perception of similarities between course material and other unrelated material and self-reports on playful approaches that questions pertaining to a nonbalance in favor of assimilation were based. A complete description of the instrumentation of the Balance concept will be discussed below.

A second variable suspected of accounting for variance in graduate level statistics scores is M-space. This is the maximum number of figurative or operational schemes that a person's short-term memory may hold at one time. Pascual-Leone (1970) suggests that problem-solving ability depends on this span of short-term memory, because it is in short-term memory that reciprocal assimilation of schemes occurs. The complexity of one's conceptual schemes is determined by the way in which they are constructed and directly by M-space. Case (1975), however, suggests that with the appropriate teaching strategy all that is required is an M-space equal two. This is because a concept can be analyzed into component concepts, and sequentially and pairwise be brought together into the superordinate concept. In the absence of this kind of instruction, M-space would be expected to be a deciding factor in course outcome.

A third variable of interest is spatial ability. Spatial visualization is defined by McGee (1979) as the "ability to mentally manipulate, rotate, twist, or invert a pictorially presented stimulus object" (p. 893). Smith (1964) suggests that spatial ability is required to make configural transformations in two and three dimensions, but also suggests that it is related to reasoning ability and the ability to grasp the Gestalt of a situation; i.e., think holistically. Piaget and Inhelder (1971) suggest that one's mental imagery is contingent upon one's operational level and, hence, one's abilities for problem-solving in general. In this sense the static, reproductive image is at the lower end of development, while the transformational, anticipatory image is at the higher end of development. This point of view was substantiated by Forisha's (1975) study in which verbal and imagery abilities were shown to develop in parallel and as a function of developmental level. A spatial score is used therefore as an inferential measure of one's level of holistic thinking and problem-solving ability in general.

Other variables were also recorded: sex and number of years of college mathematics experience. While these may be good predictors of success (Jacklin & Maccoby, 1972; deWolf, 1977), they provide little understanding of underlying processes of learning.

Of special interest to this investigation is the possible interaction of variables. It is logically expected that those with higher spatial and M-space scores do well regardless of Balance (i.e., balance of work vs. play or accommodation vs. assimilation). The differences occur under deficiencies of spatial and/or M-space scores. Logically it is expected that students deficient in both abilities would do better in a work or accommodative orientation to a complex topic such as statistics, because they lack the structures for easy assimilation,

Method

Thirty graduate students enrolled in a graduate level statistics course for behavioral sciences received course credit for completion of the following instruments:

Paper Folding Part 1 (French, Ekstrom & Price, 1963) is a test of spatial visualization involving the selection from a set of distractors the picture of an unfolded square piece of paper that had previously been folded and had a hole punched in it. The test appears to require at the highest level transformational, anticipatory imagery.

Method

Thirty graduate students enrolled in a graduate level statistics course for behavioral sciences received course credit for completion of the following instruments:

Paper Folding Part 1 (French, Ekstrom & Price, 1963) is a test of spatial visualization involving the selection from a set of distractors the picture of an unfolded square piece of paper that had been previously folded and had a hole punched in it. The test appears to require at the highest level transformational, anticipatory imagery.

Space Relations (Differential Aptitude Test Psychological Corp.) is a test of spatial visualization and surface development. One selects from a set of distractors which representation is possibly the result of folding together a cut-out and rotating it in space. The test appears to require the ability to form a transformational, anticipatory image.

A spatial score was computed by standardizing the sum of standardized scores of Paper Folding and Space Relations.

Backwards Digit Span (Case, 1977). Subjects listened to a tape of numbers being read off. Starting with three digits, going up to six digits, subjects were required to reverse the order in which the numbers were read and then write the same reversed set of digits. Subjects were monitored to make sure instructions were followed properly. Subjects had three chances for each digit span. M-space was considered to be that highest digit span for which the subject obtained at least one correct reversal.

Balance Questionnaire contained five questions pertaining to the degree with which students perceived themselves as tending to memorize statistics and their opinion as to whether statistics was "real work" (accommodative or work nonbalance). There were also five questions pertaining to the degree to which students perceive similarities between statistics and other nonmathematical material and as well whether statistics was more play than work (assimilative or play nonbalance). The final form of this questionnaire was the result of a factor analysis of an earlier form. The original questionnaire contained (as interpreted from this theoretical reference frame) three factors after rotation. These were accommodation-work, assimilation-play, and affect towards statistics. Questions loading solely on affect were not used in the final form of the questionnaire. The inter-item reliability (Brown, 1976) for the five accommodation questions was 0.85, and for the five assimilation questions was 0.75. Each item was scored 0 for not agree, 1 for agree, and 2 for strongly agree. The difference between the sum of the assimilation questions and the sum of the accommodation questions was the Balance score. Typical examples are as follows:

> It seems that I have to repeat an idea in statistics many times before I can even begin to understand it.

> Statistical ideas are so similar to some non-mathematical things I already know, I feel as if I already understand statistics because of this similarity.

The total points received in the statistics course (multiple regression) for the behavioral sciences served as criterion. Majority of exam points went to application and computation type questions.

Results

Zero-order correlations were computed among the following variables: Statistics total points, sex of the student, years of mathematics during college experience, spatial score, M-space and Balance. Years of mathematics did have a significant zero-order correlation with the criterion ($r=0.436$, $p<0.05$). However, once spatial ability, M-space, and Balance were partialled out, years of mathematics had no independent contribution to variance in the criterion ($F=1.78$, df=1,21, $p>0.15$).

TABLE 1

INDEPENDENT SOURCES OF VARIATION
IN COURSE SCORES

Source	Independent SS[a]	MS	F(1,24)	Probability
Spatial score	3.286	3.286	6.61	0.017
M-space	3.155	3.155	6.35	0.019
Balance	4.512	4.512	9.08	0.006
Spatial*M-space	5.028	5.028	10.11	0.004
Spatial*M-space*Balance	3.583	3.583	7.21	0.013
Residual (Full Model)	11.931	0.497		
Total	29.000			

[a]The change in residual sum of squares upon removal of the one source from the full model.

A statistical model was constructed by way of multiple regression (Kerlinger & Pedhazur, 1973) to account for variance in the criterion. All scores were converted to standard z-scores. In Table 1 are listed the sources of independent contributions, independent (Type IV) sums of squares, F-ratios and levels of significance (probabilities of Type I error). All main effects of interest reached levels of statistical significance ($p<0.02$). A spatial by Balance and an M-space by Balance interaction both were nonsignificant ($F=0.28$ and $F=0.19$ respectively, df=1,21, and $p>0.60$) and subsequently were pooled with the residual. Two important interactions revealed themselves through this analysis. The first is the spatial by M-space double interaction, and the spatial by M-space by Balance triple interaction. A χ^2 test for homogeneity of sample does not substantiate a possible hypothesis of a departure from sample homogeneity ($\chi^2(3)=4.003$, $p>0.20$). The regression model ($R^2=0.589$, df=5,24, $F=6.87$, $p<0.0004$) is as follows:

$$\text{Statistics} = 0.4*\text{Spatial} + 0.4*\text{M-space} + 0.4*\text{Balance} + 0.5*\text{Spatial}*\text{M-space} - 0.5*\text{Spatial}*\text{M-space}*\text{Balance}.$$

This equation is plotted as three, two-dimensional manifolds of a four-dimensional hypersurface in Figures la, b and c. Figure la is a section of the hypersurface at

which Balance =-1 standard deviations from the mean; i.e., a predominantly work or accommodative nonbalance. Figure 1b is a section in which Balance equals zero; i.e., assimilation and accommodation are in mean balance as determined by the sample. Figure 1c is a section at which Balance = +1 standard deviations above the mean; i.e., a predominantly playful or assimilation nonbalance.

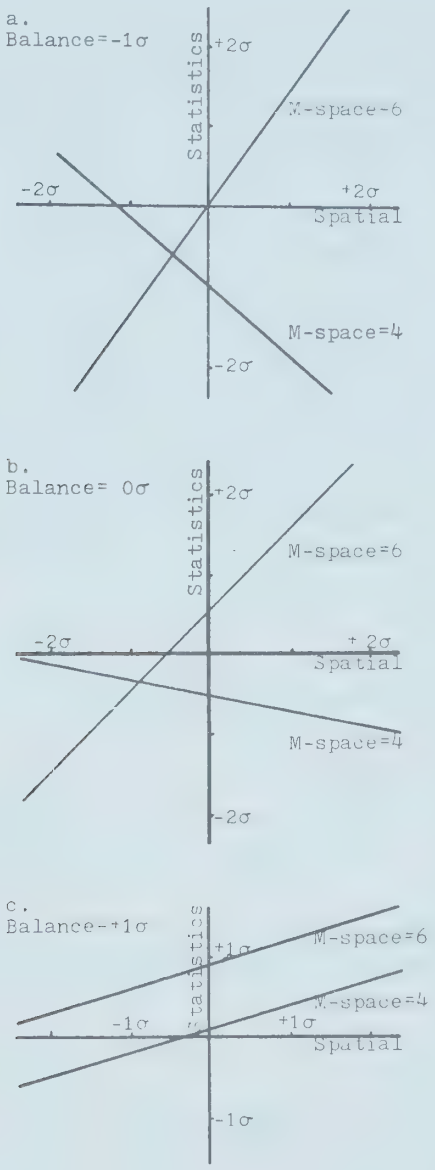

Figure 1. Three two-dimensional sections showing a Balance by M-space by Spatial interaction.

It should be noted that for spatial>0 and M-space=6, statistics course work remains high as expected across the three values of Balance. For spatial>0 and M-space=4, statistics course work increases with an increasingly assimilative or playful (Balance>0) orientation. This same trend occurs for spatial<0 and M-space=6. However, for spatial<0 and M-space=4, statistics course work increases with an increasingly accommodative or work (Balance<0) orientation. The mean M-space value for the sample was 5.13, standard deviation equals 0.86.

Discussion

It is apparent from the data that, in main effect terms, students tend to do better depending on the spatial score, M-space score, and their general nonbalance

in favor of assimilative or playful orientation. Piaget (1977) does suggest that in order for a conceptual scheme to come into "increasing equilibration" or for an accommodation to succeed, the initial nonbalance occurs by the "assimilation at the beginning of the affirmations almost exclusively" (p. 189). That is, a conceptual scheme undergoes permanent modification as a re-balancing after the similarities between the external material and the unmodified conceptual scheme have been assimilated into that same scheme. It is with respect to this theoretical point of view that the Balance main effect is interpreted. The spatial and the M-space main effects are expected by the theories discussed previously.

The disordinal interaction between spatial and M-space scores on course outcome for Balance less than or equal to mean value is complex. The regression surfaces would predict that among students with less than average spatial scores and less than positive Balance, those with a below-average M-space generally do as well as or better than others. This result is implied perhaps by Case's (1975) idea that with a proper instructional method the minimum M-space required is two. In the present situation, it would have to be the students providing their own instructional method through their learning strategies outside of class. In fact, five out of six students in this category reported taking a lot of notes so that they could learn the material out of class. It is apparent that these students did not do their learning while in class. A comparison of Figures 1a and 1b indicates that students with deficiencies both in spatial and M-space scores do better with an accommodative or work orientation than with an assimilative or playful orientation to this subject. While they do better with an accommodative orientation there is no evidence to suggest that the information is retained beyond the final exam, or in Piagetian terms, whether the conceptual schemes undergo a successful accommodation of "increasing equilibration."

For ease of interpretation, the two interaction terms uncovered through this investigation can be combined into one expression. These two terms, 0.5*Spatial*M-space and -0.5*Spatial *M-space*Balance can be factorized as follows:

0.5*(1 - Balance)*Spatial*M-space.

It is then interpreted that Balance affects the Spatial by M-space interaction via its regression weight.

References

Brown, F. G. *Principles of educational and psychological testing.* New York: Holt, Rinehart & Winston, 1976.

Case, R. Learning and development: A neo-Piagetian interpretation. *Human Development*, 1972, *15*, 339-358.

Case, R. Structures and strictures: Some functional limits on the course of cognitive growth. *Cognitive Psychology*, 1974, *6*, 544-573.

Case, R. Gearing the demands of instruction to the developmental capacities of the learner. *Review of Educational Research*, 1975, *45*(1), 59-87.

Case, R. Responsiveness to conservation training as a function of induced subjective uncertainty, M-space, and cognitive style. *Canadian Journal of Behavioral Science*, 1977, *9*(1), 12-25.

deWolf, V. A. High school mathematics preparation and sex differences in quantitative abilities. Educational Assessment Center Report, 1977, No. EAC-R-77-17 Washington University, Seattle, Washington.

Forisha, B.D. Mental imagery, verbal processes: A developmental study. *Developmental Psychology*, 1975, *11*(3), 259-267.

French, J., Ekstrom, R., & Price L. *Kit of reference tests for cognitive factors*, Princeton, New Jersey Educational Testing Service, 1963.

Ginsburg, H. and Opper, S. *Piaget 's theory of intellectual development*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1979.

Jacklin, C.N., & Maccoby, E.E. Sex differences in intellectual abilities. Paper presented at AERA 1972 No. ED063545 (ERIC).

Kerlinger, F. N., & Pedhazur, E. J. *Multiple regression in behavioral research*. New York: Holt, Rinehart and Winston, 1973.

McGee, M. G. Human spatial abilities: Psychometric studies and environment, genetic, hormonal, and neurological influences. *Psychology Bulletin*, 1979, *86*(5), 889-918.

Pascual-Leone, J. A mathematical model for the transition rule in Piaget's developmental stages. *Acta Psychologia*, 1970, *32*, 301-345.

Phillips, J. L. *The origins of intellect: Piaget's theory*, San Francisco: W.H. Freeman & Co., 1975.

Piaget, J. Science education and the child. New York: Basic Books, 1970.

Piaget, J. *The development of thought*. New York: Viking Press, 1977.

Piaget, J., & Inhelder, B. *Mental imagery and the child*. New York: Basic Books, Inc., 1971.

Samples, R. E. Notes on teaching science. *Science and Children*, 1968, *5*(6), 28-30.

Samples, R. E. *The metaphoric mind*. Reading, Massachusetts: Addison, 1976.

Smith, I. MacFarlane. *Spatial ability*. London: University, 1964.

The Alberta Journal of Educational Research *Vol. XXVIII, No. 2, June, 1982, 122-134*

Y. L. JACK LAM
Brandon University

Teacher Professional Profile: A Personal and Contextual Analysis

The professionalization of teachers, with its profound impact upon the power structure in the Educational System and program delivery, has received much discussion but relatively few empirical investigations. Using Snizek's shorter version of Hall's professionalism scale assessing five attitudinal components, provincial professional profiles have been developed from 1204 Manitoba teachers based on six personal and contextual variables. Results indicated that with the exception of the "perceived usefulness of professional development," no clear-cut patterns can be established between professionalism and variables selected, casting doubts on many of the common assumptions and generalizations.

Contributing to the basic changes in the educational system in the past three decades has been a movement of professionalization undertaken by teachers and their organizations. Profusely and casually used, the term "professionalization" in the context of public school education has been rarely clarified or understood. Subsequent debate and reaction has tended to underscore the significant ramifications of this issue for the present and future educational governance structure and program delivery.

Resolution to the current controversy rests on a careful re-examination of the classical definition of professionalization provided by Vollmer and Mills (1966, pp. vii-viii). They referred to the term as "a dynamic process whereby many occupations can be observed to change certain crucial characteristics in the direction of a 'profession'." Thus a static comparison of any occupation in the process of change, like teaching, against an ideal set of criteria (e.g. Becker, 1963; Eisenstein, 1972; McDaniel, 1979) is far less meaningful than placing the occupations concerned along a professional continuum. Interpretation would then be focused on the degree of professionalism one occupation displays at a given point in time and space.

Y. L. J. Lam, (PhD, University of Toronto) is Associate Professor and Chairman of the Department of Administration and Educational Services, Faculty of Education, Brandon University. The focus of his research is teacher professionalism, school organizational environmental models, and adult education.

Imbedded in the definition too are two main aspects of professionalization, structural and attitudinal, which, left undemarcated, account for much of the current confusion. The structural aspect of professionalization pertains to the authority structure governing the behavioral and ethical standards of members of the occupation as well as to the recognition endorsed by laws and society. To the extent that the loci of the power are still located in the provincial Department or Ministry of Education, the school board, and central office staff, and to the extent that society has not yet equated teaching to other established professions like medicine and law, and that some legislative act has yet to be engraved for treating teaching as a profession, one has to concede that professional self-designations by teacher organizations do not constitute criteria of a professionalized occupation.

Historically, barriers to structural changes in teaching originated from a number of sources. One was associated with the common conviction of pluralistic control of education which defies monopoly by members of teacher organizations. Another was attributed to the domination of the occupation by married women whose traditional commitments to families and careers were in constant conflict and mutual hindrance. Still another was related to the narrowness of the expertise gap between teachers and laymen which creates problems for securing instant public respect (Cox & Elmore, 1976).

Militancy of teachers in recent years has been considered by many (e.g., Corwin, 1965; Cox & Elmore, 1976) as a natural development of the professionalization process which attempts to restructure the governance system in public education. While growing activism has made some headway in improving teacher status, it is perhaps too early to pass judgement on how successful teacher organizations are in attaining structural changes more in line with the professional characteristics.

Perhaps more critical to an understanding of the professionalization of teaching is the study of the attitudinal dimension. Arguing from the premise that attitude governs one's receptivity to and acceptance of certain kinds of theories and knowledge, which in turn affects the skills or behaviors one displays (Lam, 1979a), the examination of attitudes of teachers unveils the mentality, psychological orientation, and morale which are intricately related to program delivery and quality of instruction. Through an extensive search of the literature, Hall (1967) identified five major aspects which constitute the conceptual basis of his professionalism scale. These are

1. use of the professional organization as a major referent
2. belief in public service
3. belief in self-regulation
4. sense of calling to the field, and
5. autonomy.

Snizek (1972), by additional analysis, had verified 25 of the original 50 items that empirically best correspond to these theoretical dimensions.

The primary purpose of this paper is to document the application of the revised Hall's instrument to the assessment of teachers' attitudes of professionalism. Following McDiarmid's assertion (1970) that some portions of the teacher population may be considered more professional than others, and Hill's observation (1975) that there is a wide range of levels of commitment to educational values exhibited by members of the occupation, the specific objective here is to construct teachers' profiles of professionalism on the basis of selected individual and contextual attributes discussed in literature. Included in the

present study are sex, teaching experience, grade level taught, regional differences, formal education, and professional development.

That sex is often considered a deterrent to teachers' professionalism is argued for three reasons: First, the majority of the married female teachers have yet to resolve the role conflict of family and career commitments. Second, in the continuing trend of bureaucratization of the education system, women teachers were less alienated than their male counterparts (e.g., Hearn, 1971). Third, women teachers' preference for non-aggressive behaviors (Prather, 1971) contributes little to the militant mentality so vital for the professionalization process. In contrast, the increasing proportion of males entering teaching was emphatically linked to the rise of activism (Hennessy, 1975; Selinger & Goldhammer, 1972). Such an assertion, however, was not fully endorsed. Nastrom and Butler (1975) reported that women exhibited a higher degree of professionalism than did men; no difference in professionalism between women teachers of single or married status was found. Additional investigation in this area is warranted.

Teaching experience is another attribute often related to professionalism. In studying teachers' career commitment and occupational attitudes, Lauglo (1975) found older and experienced teachers had a stronger pedagogical commitment during their university years, were more satisfied with their occupational status, and had a stronger present commitment to their profession. Experience was also found to be associated with greater innovation in classroom instruction and greater assertion of professional autonomy in different cultural settings (Lam, 1975).

Grade level was another contextual variable affecting the attitude of teachers. While examining teachers' reactions to decision-making roles and responsibilities, Talmange and Orstein (1973) reported that teachers of senior grades were more anxious to be involved. Cortis (1973) also found that in contrast to elementary teachers, secondary teachers were more sensitive yet more tolerant in personality. They possessed a more "progressive" attitude and expressed a higher degree of satisfaction with teaching.

Regional differences, as independent contextual attributes, have seldom been examined closely. However, in view of community-related lifestyles of teachers and different school characteristics, one would anticipate urban-rural multifactor analysis of patterns of teacher participation in school system decision-making. Alutta and Belasco (1973) showed that young males in secondary *rural* areas perceived the highest level of decisional deprivation while older females in elementary *urban* areas perceived the greatest degree of decisional saturation. Much indeed awaits further exploration to illuminate attitudinal differences based on regions.

Level of formal education was another factor associated with professionalism. Taking a historical perspective, Bullough (1970) stressed the institutional and social importance of the university to the concept of professionalization and argued forcefully that a professional image is forged by the level of formal education attained. Empirically, Pavalko (1970) indicated that teachers who did graduate work were more committed to teaching careers than those who did not. Cortis (1973) also detected a link between professional preparation at college and teaching abilities.

Professional development, encompassing all types of inservice activities, is designed implicitly and explicitly for teachers to grow on the job and should have

significant impact on the professionalism of teachers. However, empirical evidence is dubious about its long lasting effects on instruction (Lam, 1979b), and much has yet to be learned regarding its impact on professional attitudes.

Research Design

In order to develop profiles representative of the entire province of Manitoba, stratified random sampling of school divisions on the basis of location was resorted to. All schools in the selected divisions were approached to cooperate. Subsequently, 1204 teachers participated, representing one-tenth of the total teacher population in the province from 114 schools in seven divisions situated in large metro, small urban, as well as medium and small rural settings.

Each teacher was asked to respond to the 25 items that Snizek (1972) identified as best from Hall's professionalism scale. These 25 items, with five measuring each dimension, were accompanied by five-point Likert-type scales ranging from "very well" to "very poorly." In addition, data on the six individual and context variables were also collected.

Interrelationship among personal and contextual variables were assessed by correlation analysis to determine the degree of contamination effects that should be borne in mind while interpreting findings. Cross-tabulation by subgroups in each personal and contextual variable provided the scoring pattern as well as the total and mean score for each group for each professional subscale. Summation and averaging of the five subscales then provide the composite index for each professional dimension to be used for profile construction and statistical analyses.

Multitudes of ONE-WAY Analyses of Variance were then employed comparing subgroups of each personal and contextual variable on each dimension of professionalism to determine areas of statistical differences.

Analysis and Discussion

From the zero-order correlation analysis (Table 1) which probed interrelationships among personal and contextual variables, it was revealed that setting was related negatively to formal education, a suggestion that teachers from metro divisions were more educated than those from small urban and rural areas. Also longer periods of teaching service were found among secondary teachers and it was also this group that benefited most from professional development activities.

Professional Profiles by Sex

The profiles compiled from 701 men teachers and 503 women teachers (Figure 1) were an indication that men teachers were somewhat higher than women teachers on "using the professional organization as a major referent," "belief in public service," and "sense of calling to the field." On the other hand, women teachers scored slightly higher in "belief in self-regulation" and "autonomy." Upon closer examination of results of Analysis of Variance (Table 2), it was indicated that male teachers scored significantly higher on "use of the professional organization as a major referent" whereas female teachers were significantly higher on "sense of calling to the field."

These results do not endorse the widely held assumption that female teachers are less professional than male teachers. Neither do they provide a blanket statement to support that women exhibited a higher degree of professionalism

than did men. To the extent that men teachers showed greater interest and faith in their organization and to the extent that women teachers were more confident of competence and devotion to their jobs, one could draw the tentative conclusion that sex made a difference in professional orientation: Men teachers' concern for the structural aspects of professionalism probably induced them to be more involved with organization activities. This supports the observation that an increasing proportion of males entering teaching contributed to the growing militancy in recent decades (e.g., Selinger & Goldhammer, 1972; Hennessy, 1975).

TABLE 1

ZERO-ORDER CORRELATION ANALYSIS
OF PERSONAL AND CONTEXTUAL VARIABLES

Variables	Setting	Sex	Grade Level	Teaching Experience	Formal Education
Sex	.01				
Grade Level	.00	.15			
Teaching Experience	-.02	.01	.31*		
Formal Education	-.27*	.18	.16	.06	
Professional Development	.01	-.08	.30*	.04	.08

* Significant at .01 or better.

Female teachers, on the other hand, are more concerned about the execution of classroom responsibilities and about teaching as a career. Social and economic conditions of the late 70's must have transformed women teachers' in-and-out career pattern of the 60's to one of greater permanency.

Professional Profiles by Experience

Professional profiles generated from 289 teachers having one to five years of experience, 288 whose services ranged from six to ten years, and 627 who had spent more than ten years in the school system revealed that no simple direct relationship between experience and professionalism could be detected. Composite score for the least experienced was slightly higher than other groups in Dimension II, i.e. "belief in public services." The most experienced group, while scoring higher in several aspects, was lowest in the dimension of "autonomy," an aspect in which teachers with a medium range of experience outscored other groups. Additional analysis of variance revealed that differences in teaching experiences accounted for significant variation in four of five professional dimensions. In detail, those with more than ten years of teaching experience scored highest in "belief in public service." On the other hand, those with the least experience excelled in "belief in self-regulation," "sense of calling to the field," and even "autonomy."

Apparently, longevity of service cultivates a sense of pride in service to the community. However, it is the enthusiasm characteristic of the least experienced

TABLE 2

SUMMARY OF ONE-WAY ANOVA ANALYSES OF PERSONAL AND CONTEXTUAL FACTORS' EFFECTS ON TEACHER PROFESSIONALISM

	DIMENSIONS OF TEACHER PROFESSIONALISM																								
	I					II					III					IV					V				
	$\bar{X}$	S.D.	DF	S.S.	F	$\bar{X}$	S.D.	DF	S.S.	FF	$\bar{X}$	S.D.	DF	S.S.	F	$\bar{X}$	S.D.	DF	S.S.	F	$\bar{X}$	S.D.	DF	S.S.	F
Sex																									
1. Male	14.30	1.85		14.77		7.28	1.42		.11		4.27	1.37		.65		7.30	1.81		16.91		5.34	1.47		3.18	
			1		4.05*			1		.06			1		.35			1		5.05*			1		1.42
2. Female	14.08	1.98		4387.60		7.26	1.31		2301.55		4.22	1.36		2257.93		7.54	1.83		4019.17		5.24	1.52		2687.41	
Teaching Experience																									
1. 1-5 years	14.21	1.84		10.13		7.16	1.35		21.98		4.49	1.39		22.46		7.30	1.75		37.47		5.56	1.39		27.94	
			2		1.38			2		5.77**			2		6.00**			2		5.60**			2		6.30**
2. 6-10 years	14.36	1.89		4319.03		7.10	1.39		2252.28		4.14	1.34		2214.54		7.06	1.82		3955.94		5.19	1.53		2624.63	
3. Over 10 yrs.	14.13	1.94				7.40	1.38				4.25	1.39				7.20	1.91				5.23	1.47			
Grade Level																									
1. Elementary	14.45	1.79		3.45		7.57	1.37		15.22		4.18	1.25		.31		7.47	1.55		18.56		5.61	1.34		10.82	
			1		.94			1		4.73*			1		.16			1		5.54*			1		4.85*
2. Secondary	14.20	1.91		4398.95		7.26	1.38		2296.48		4.25	1.37		2258.27		7.14	1.84		4036.14		5.28	1.50		2684.71	
Setting																									
1. Metro	14.16	1.95		19.18		7.24	1.41		2.36		4.38	1.38		34.45		7.34	1.88		38.97		5.38	1.55		84.42	
2. Small Urban	14.11	1.85				7.32	1.43				4.08	1.37				7.41	1.82				4.74	1.32			
			2		2.63			2		.62			2		9.30**			2		5.85**			2		19.45**
3. Rural	14.46	1.80		4383.20		7.34	1.24		2299.32		3.99	1.25		2214.15		7.80	1.61		3997.19		5.53	1.33		2606.22	
Formal Education																									
1.-	13.93	1.73				7.84	1.35				4.50	1.50				8.11	2.14				5.52	1.49			
				51.02					27.17					12.24					38.50					22.16	
2. Years of	13.77	1.73				7.35	1.77				4.22	1.39				7.63	1.98				5.61	1.43			
			5		2.80**			5		5.43**			5		1.30			5		2.32*			5		1.97
3. Post-secondary	14.06	1.94				7.07	1.70				4.28	1.46				7.58	2.19				5.45	1.57			
4. Education	14.19	1.92		4292.38		7.26	1.29		2240.47		4.22	1.34		2218.21		7.46	1.73		3906.88		5.34	1.36		2650.20	
5.	14.21	1.97				7.28	1.37				4.31	1.38				7.34	1.70				5.21	1.56			
6.4	14.68	1.96				7.12	1.22				4.04	1.30				7.31	1.94				5.08	1.69			
Professional Development																									
1. Useless	13.50	2.62				7.31	1.66				4.74	1.48				6.78	2.07				5.89	1.68			
				175.73					.38					42.79					102.25					106.33	
2. Somewhat useful	13.97	1.85				7.29	1.36				4.32	1.33				7.31	1.78				5.46	1.39			
			2		24.99**			2		.09			2		11.57**			2		15.56**			2		24.73**
3. Useful	14.64	1.69		4169.22		7.26	1.34		2286.78		4.08	1.36		2192.77		7.74	1.78		3896.29		4.97	1.50		2549.03	

* Significant at .05
** Significant at .01 or better

teachers that might have accounted for this group to display a stronger feeling towards continued self-improvement, a greater devotion to their job, and less inhibition in carrying out what they deemed essential within their responsibilities. By examining the mean scores of each group, it is also interesting to note that with the exception of the first dimension, teachers with medium experience scored much lower than their least and most experienced counterparts. It is likely that, in terms of a teaching career, medium experience represents a "depression" period — a time when the initial enthusiasm on the job begins to wear off but the commitment to the profession has yet to materialize. Further exploration in future may shed light in this uncharted domain.

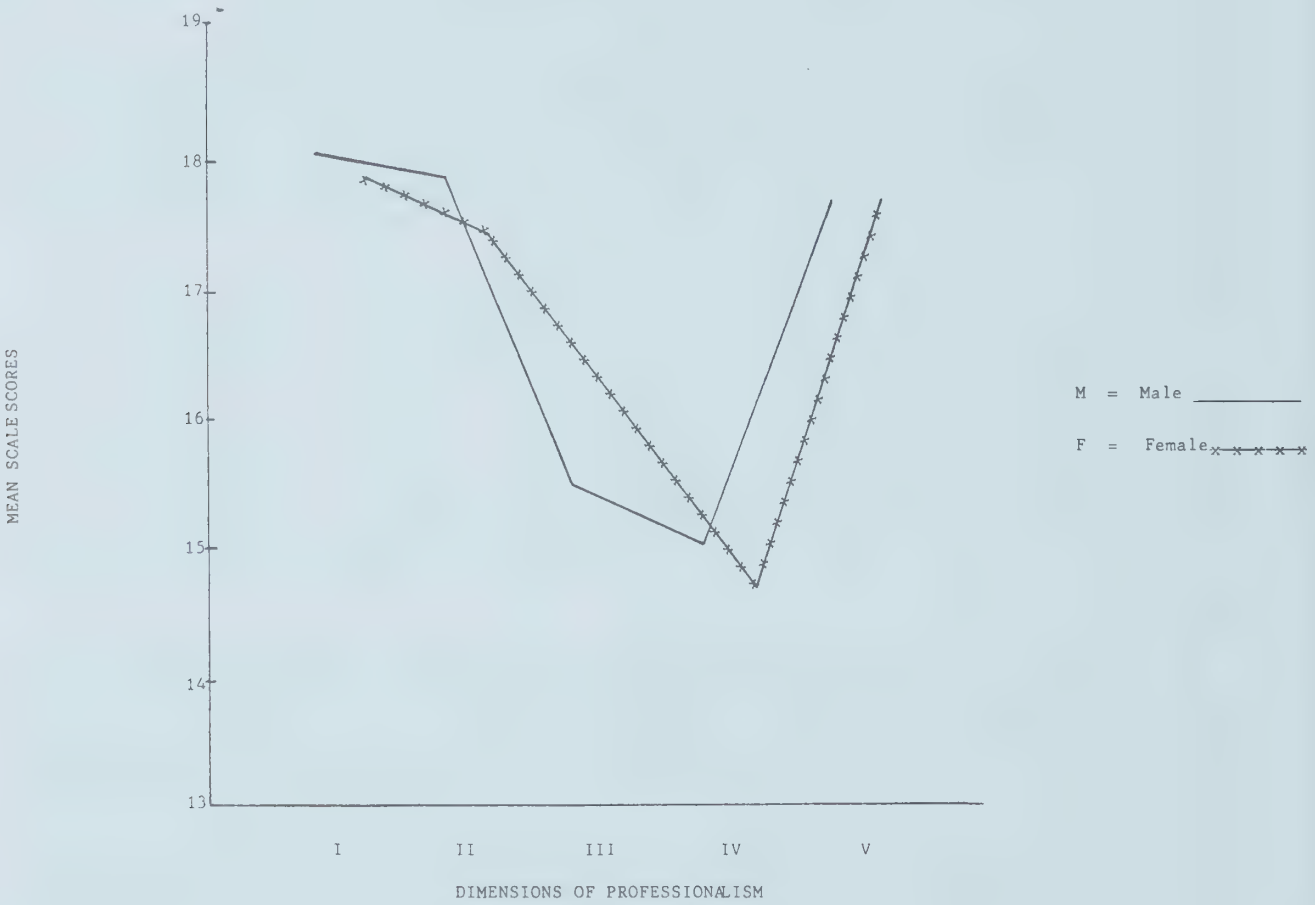

Figure 1. Profile of teacher professionalism by sex.

Professional Profiles by Grade Level Taught

Compiled from 880 elementary and 324 secondary school teachers in the sample, the profiles indicated that elementary school teachers scored slightly higher on "belief in public service," "belief in self-regulation," and "sense of calling to the field," but lower on "using the professional organization as a major referent" and "autonomy." Analyses of variance revealed three specific areas where statistical differences were registered. First, secondary school teachers were less certain about teaching as an important public service. Second, elementary school teachers seemed to express greater devotion to the job. Third, secondary school teachers felt they had greater latitude to exercise autonomy.

Notwithstanding greater decision-making power at the secondary level, attributable also to longer period of services, these teachers were involved in the

final stage of the critical formative educational process. As such, it is likely that they felt they had little part to play in shaping children's development, and this induced them to doubt the contribution of their services, more so than the elementary school teachers. Thus, even though findings of the present study do not contradict existing literature, it seems certain that enjoying greater decision-making roles and responsibilities (Talmange & Orstein, 1973) and having a higher degree of satisfaction (Cortis, 1973) do not constitute greater professionalism in all aspects.

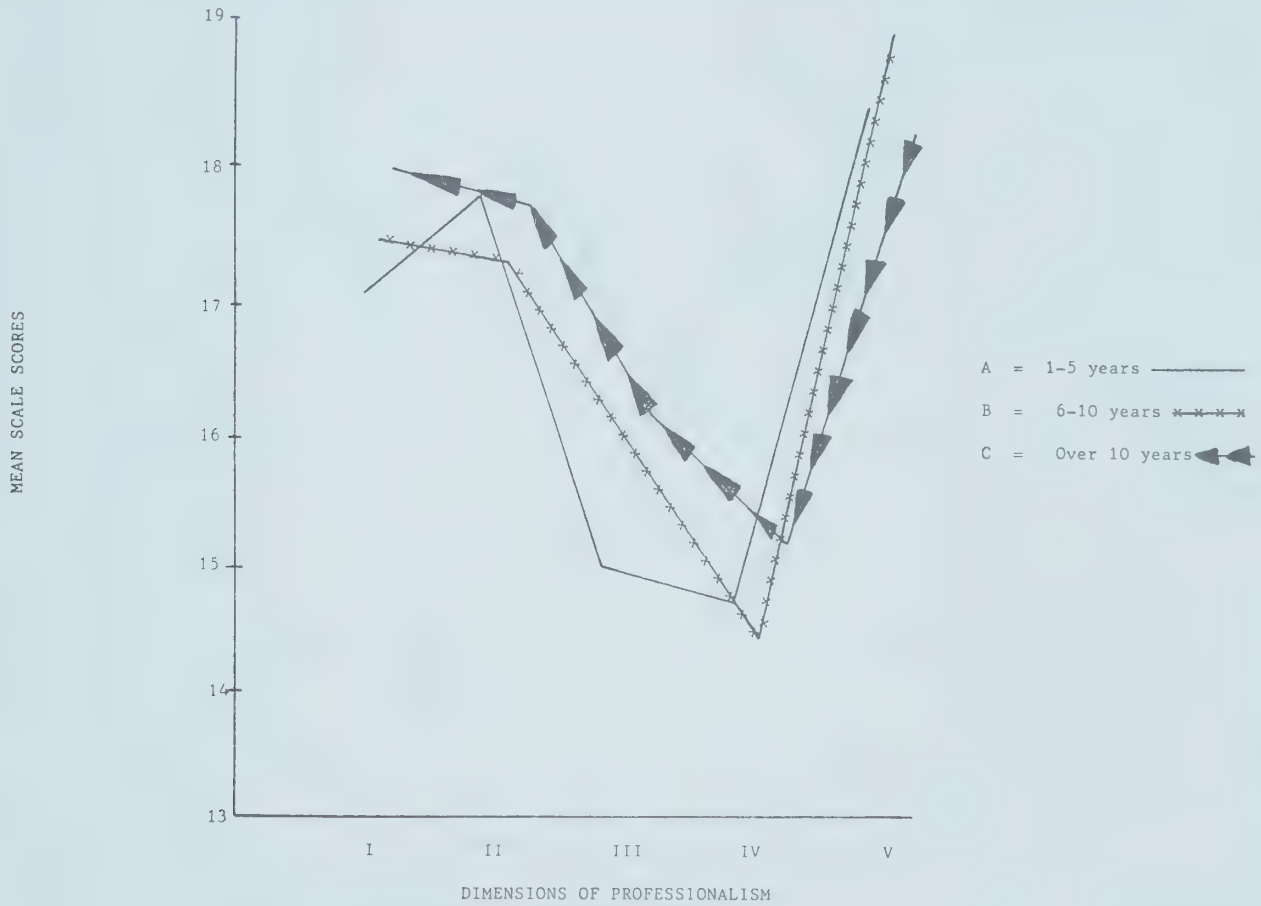

Figure 2. Profile of teacher professionalism by teaching experience.

Professional Profiles by Regions

Based on the responses of 759 teachers from the metro area, 210 from smaller urban centres in the province, and 235 teachers from the rural divisions, the profiles for the three groups did not concur with the general expectation that the more "sophisticated" teachers from the major urban areas were more professional than teachers from smaller urban and rural divisions. Indeed on two dimensions, i.e., "sense of calling to the field" and "autonomy," teachers from the rural division scored significantly higher than other groups. It was in the "Belief in self-regulation" that teachers from the metro and small urban area had higher composite scores. Regional differences in professionalism may reflect characteristics of the community as well as those of individuals. In communities where fewer demands are competing for teachers' time, as in the case of small urban or rural areas, teachers are more inclined to focus more on the job and this in turn reinforces the feeling of dedication to education. With size of school being

smaller and school structure being simpler, rural teachers feel that they have a chance to exercise greater discretion in carrying out their responsibilities, thus enhancing the sense of autonomy.

On the other hand, remoteness from university campuses or centres of postgraduate studies in the rural divisions reduces teachers' chance and, perhaps, desire for ongoing academic and professional growth, a point well illustrated by significant relationship between setting and formal education in earlier analysis. Subsequently, "belief in self-regulation" was significantly lower than that of their urban colleagues.

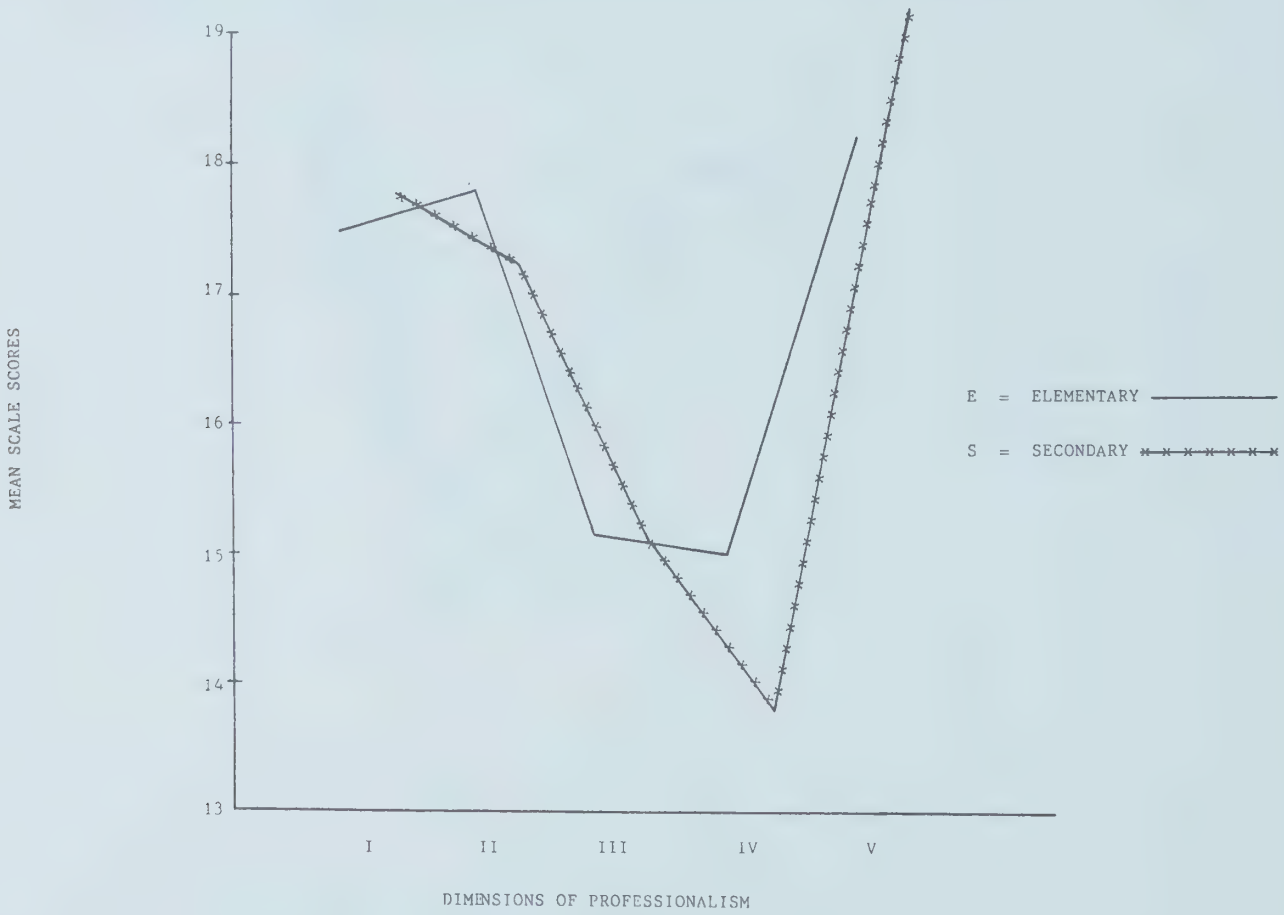

Figure 3. Profile of teacher professionalism by grade level.

Professional Profiles by Formal Education

Profiles generated from 64 teachers with one year of postsecondary training, 68 teachers with two years, 69 with three years, 441 with four years, 407 with five years, and 136 with six years of postsecondary education do not convey the impression that more formal education was conducive to greater professionalism. In two aspects, in fact, the least trained teachers scored much higher than other groups. They demonstrated a stronger belief in the importance of teacher services to the public and they showed a stronger sense of calling to the field. The most qualified teachers, on the other hand, scored somewhat higher in "belief in self-regulation," and the next best qualified teachers, in using "the professional organization as a major referent" and "autonomy." ANOVA Analyses confirmed most of the profile patterns. The best qualified teachers excelled only in their dependence and support for their professional organization. One can speculate

that more formal education promotes greater awareness of and involvement in professional activities. In these respects, the more qualified teachers are identified as "activist" or "militant" in literature. The reversed impact on belief in teaching services, devotion, and idealism might be attributed either to the increasing experiences beyond the physical boundary of the classroom or to more career alternatives available in fields other than teaching for the better qualified teachers.

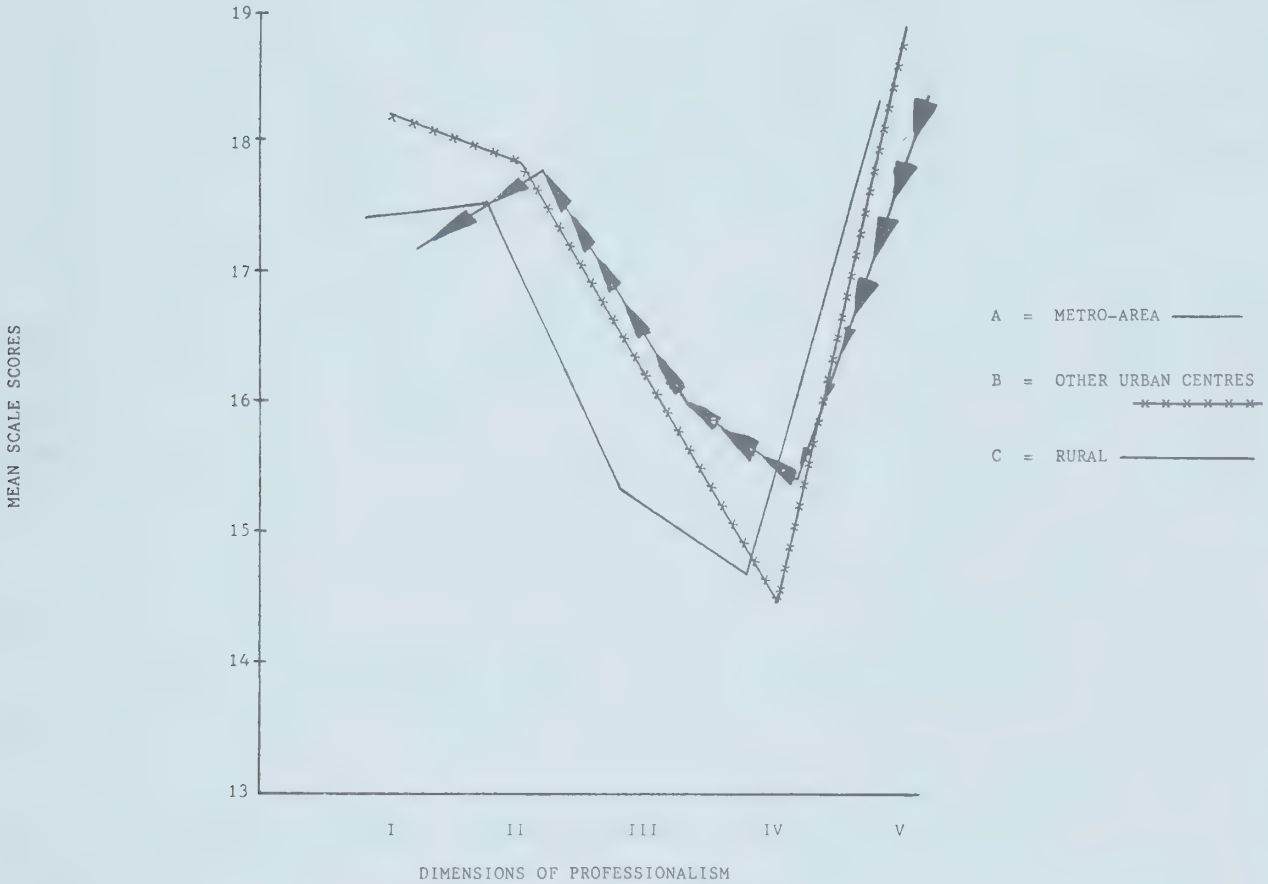

Figure 4. Profile of teacher professionalism by region.

Professional Profiles by Perceived Usefulness of Professional Development

Irrespective of the actual impact professional development activities had upon the growth of teachers on the job, the profiles displayed by teachers grouped in terms of their perception of the usefulness of these activities followed very clear-cut patterns. Those who perceived professional development to be useful had the highest scores on all aspects of professionalism, whereas those who felt professional development was not beneficial scored lowest on all five dimensions. Analyses of Variance indicated that the "benefited" groups scored statistically higher in their support for the professional organization. They also displayed a greater "sense of calling to the field" and were more inclined to exercise autonomy on the job.

It would seem then that those who benefited from the professional development activities were superior in three aspects of professionalism. While the literature continues to question the lasting impact of professional development

activities on teaching performance, there is little doubt from the present study that these inservice activities greatly enhanced the attitudinal aspect of professionalism.

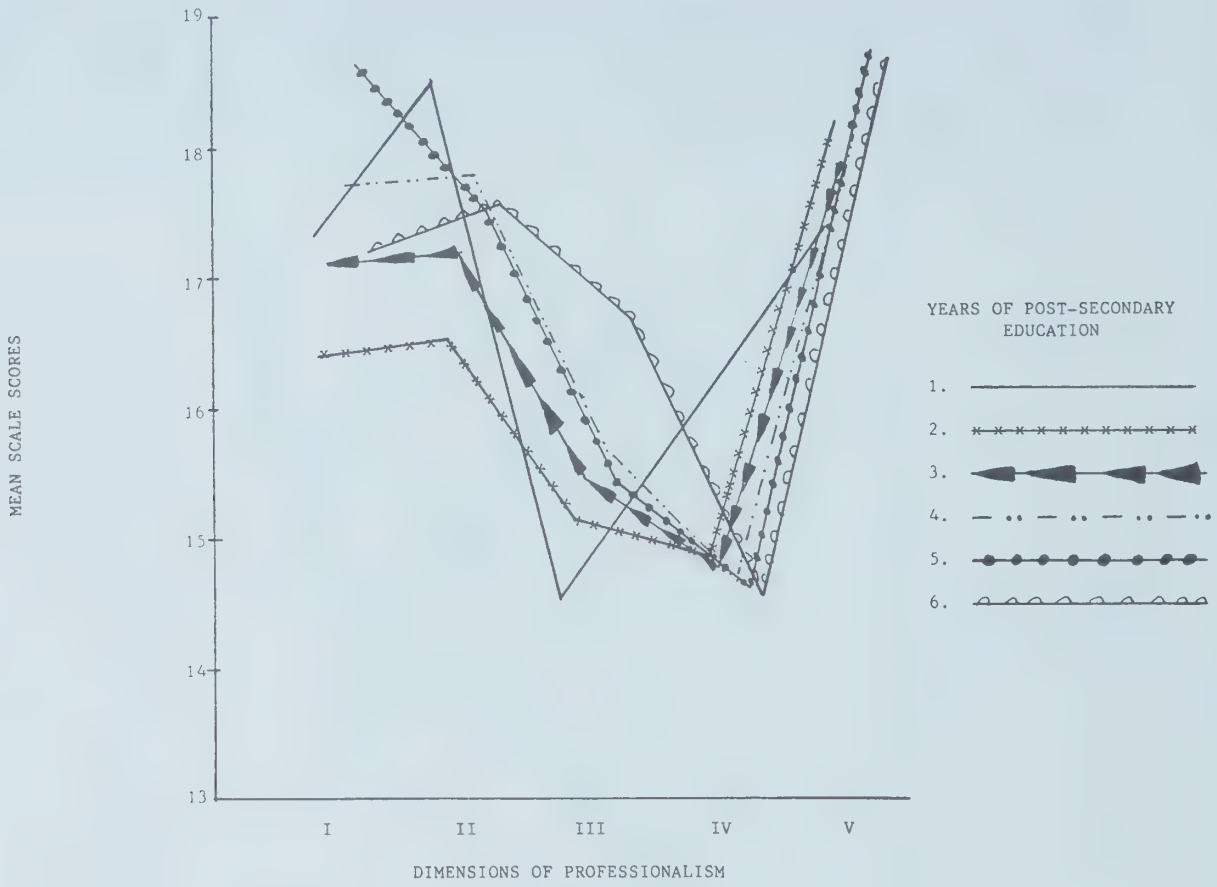

Figure 5. Profile of teacher professionalism by formal education.

Conclusion

Irrespective of substantial problems that remain unresolved in attaining the structural aspect of professionalization, from the professional profiles developed from one-tenth of the total teacher population in Manitoba, it was indicated that, on the attitudinal dimension of professionalism, teachers in general scored well above the absolute scale mean. Overall, teachers use their professional organization as a major referent, they have few doubts about the importance of teaching as a service to society, and they achieve a high degree of autonomy in carrying out their responsibilities. A great portion of them, too, have a strong belief in self-regulation. The only dimension that appears to be in doubt is the sense of calling to the field. In this one area, teachers seem to score lower than the absolute scale mean. This may reflect the general morale of a profession in a period saturated with problems and uncertainty. Declining enrolment coupled with staff layoff in the public education system, the tightening of fiscal spending, the difficulty of annual negotiations, and the public demand for accountability offer little encouragement to people who choose teaching as their professional career. Of course, further empirical analysis is required to substantiate the speculated effects of these external factors on teachers' professional attitudes.

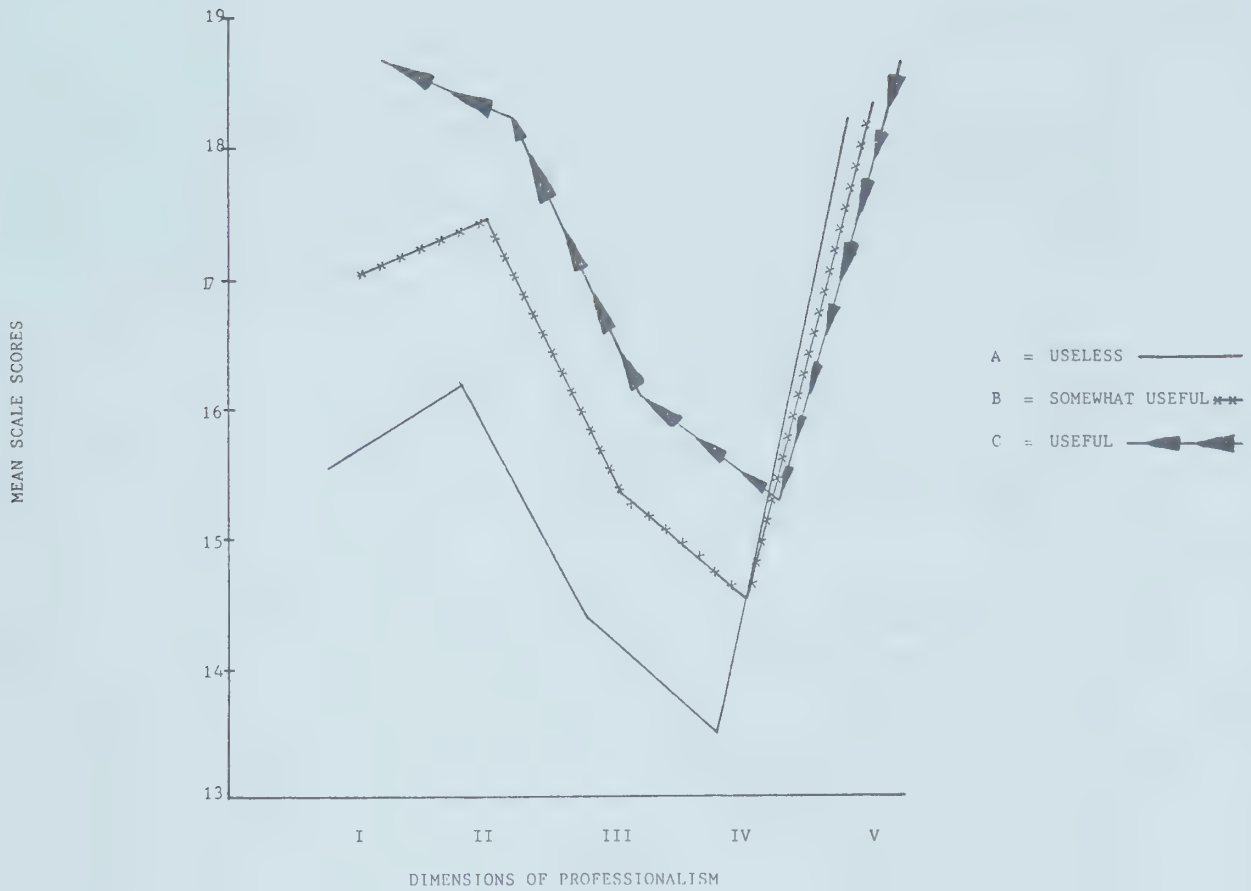

Figure 6. Profile of teacher professionalism by perceived usefulness of professional development.

In a more specific scrutiny of the professional profiles developed on personal and contextual variables, one derives an impression that the existing literature may be over-simplistic when attempting to make global statements of teachers' professionalism based on single attributes. Given the possibility of the interactive effects of the variables examined, one concludes that no simple linear relationship can be established between professionalism and most of the factors used. The defiance to our attempt to identify clear-cut patterns of professionalism reflects the complexity of reality and calls for further validation. An exception to this conclusion comes from the perceived usefulness of professional development. While the influence of the current format of inservice activities on the improvement of job performance is very much in doubt, they have proved important at least in the development of professional attitudes. Teacher organizations, school divisions, and faculties of education in the universities might continue to utilize this vehicle in promoting attitudinal professionalism, and this, in turn, may have a ripple effect on the entire process of professionalization.

Acknowledgement is made to the University Research Grant for funding and to the teachers in the seven school divisions for assistance in data collection.

References

Allutta, J. A. & Belasco, J. A. Patterns of teacher participation in school system decision making. *Educational Administration Quarterly*, 1973, *9*(1), 21-41.

Becker, H. The teacher in the authority system of the public school. *Journal of Educational Sociology*, 1963, Nov., 136-144.

Bullough, V. L. Education and professionalization: A historical example. *History of Education Quarterly*, 1970, *10*(2), 160-169.

Cortis, G. A. An analysis of some differences between primary and secondary teachers. *Educational Research*, 1973, *15*(2), 109-144.

Corwin, R. G. Militant professionalism , initiative and compliance in public education. *Sociology of Education*, 1965, *38*, 310-331.

Cox, H., & Elmore, J. Public school teachers: A sociological view of their status. *Contemporary Education*, 1976, *47*(4), 244-247.

Eisentein, H. S. Professionalization and teaching: A critiquè. *Urban Education*, 1972, *7*(1), 79-94.

Hall, R. Components of professionalization. Paper presented at Annual Meeting of the American Sociological Association, San Francisco, California, 1967.

Hearn, J. J. Teachers' sense of alienation with respect to school system structure. *Phi Delta Kappan*, 1971, *53* (5), 312.

Hennessy, P. H. Teacher militancy. Canadian Teacher Federation: Ontario, 1975, 1-80.

Hill, J. Obstacles to professional self-governance in education. *Administrative Notebook*, 1975, *24*(4), 1-4.

Lam, Y. L. J. Educational technology in different socio-cultural contexts. *Canadian and International Education*, 1975, *4*(1), 55-68.

Lam, Y. L. J. Assessing the needs of school administrators: A critical analysis. *Challenge*, 1979, *18*(4), 15022. (a)

Lam, Y. L. J. Effectiveness of an inservice workshop: A case analysis. *Manitoba Journal of Education*, 1979, *14*(1), 25-32. (b)

Lauglo, J. Teachers' social origins: Career commitment during university and occupational attitudes. *Sociology of Education*, 1975, *48*(3), 287-307.

McDaniel, T. R. The de-professionalization of teachers. *Educational Forum*, 1979 *43*(2), 229-237.

McDiarmid, G. L. The myth of professionalism. *O.S.S.R.F. Bulletin*, 1970, *50*(1), 17-20.

Nastrom, R. R., & Butler, W. E. The professionalism of women teachers. *Kappa Delta Phi*, 1975, *12*(1), 6-8.

Pavalko, R. M. Recruitment to teaching: Patterns of selection and retention. *Sociology of Education*, 1970, *43*(3), 340-353.

Prather, J. Why can't women be more like men: A summary of the sociopsychological factors hindering women's advancement in the professions. *American Behavioral Scientist*, 1971, *15*(2), 172-182.

Selinger, A., & Goldhammer, K. The roots of teacher militancy. In P. J. Cistone (Ed.), *School boards and the political fact*. Toronto: O. I. S. E., 1972, 45-52.

Snizek, W. E. Hall's professionalism scale: An empirical reassessment. *American Sociological Review*, 1972, *37*, 109-114.

Talmange, H., & Orstein, A. C. Teachers' perception of decision making in defining accountability. *Journal of Negro Education*, 1973, *8*(2), 212-221.

Vollmer, H. M., & Mills, D. L. (Eds.). *Professionalization*. Englewood Cliffs, New Jersey: Prentice-Hall Inc., 1966.

The Alberta Journal of Educational Research *Vol. XXVIII, No. 2, June, 1982, 135-148*

S. M. SHAPSON
and
E. M. DAY
Simon Fraser University

A Comparison Study of Three Late Immersion Programs

In this study, a comparison is made between the effects of three grade 6 late immersion programs in British Columbia which differed in the amount of instructional time in French (90% vs. 60% in grade 6), the previous background of the students in French, and the use of selection criteria. The results of comparisons at the beginning of the program revealed (a) no significant differences among the students in the three programs in cognitive abilities and (b) significant differences in French language skills and attitudes toward French language and culture. From the results of comparisons at the end of grade 6, significant differences were shown in attitudes but not in French comprehension or achievement in English, mathematics, science, or social studies. No significant differences in French comprehension or French language arts were revealed in follow-up comparisons at the end of grade 7, which may be an indication that a late immersion program can be equally effective under a variety of conditions.

The past fifteen years have seen a great increase in the number and diversity of educational opportunities for English-speaking Canadian children to learn French as a second language. This increase has been prompted by the desire of many Canadians for the improvement and strengthening of French language instruction, and it has been stimulated by funding from the Federal Government, which has encouraged innovation and research in second language programs. Of

---

Stanley M. Shapson (PhD, Developmental Psychology, York University) is Associate Professor and Director of Professional Programs in the Faculty of Education, Simon Fraser University. His research interests include Bilingual/Multicultural Education, and Program Evaluation and Policy Formation of Educational Programs. He has been engaged in several research studies funded by governments, including studies about elementary French programs in B.C., national immersion program for postsecondary students, and effects of class size in junior grades.

Elaine Day is a Research Associate with the B.C. French Study research team, Faculty of Education, Simon Fraser University. Her main interests are in psycholinguistics and second language acquisition. She has collaborated in several studies of early and late immersion programs.

the variety of programs which have resulted, immersion appears to be among the most effective for enabling English-speaking children to attain high levels of proficiency in French. The immersion program typically involves majority language speakers who are learning a useful, relatively prestigious language; thus its applicability should not be generalized to minority language situations (see Cohen, 1976, for a discussion of the differences between immersion programs and language programs for minority language children).

In an immersion program, all or a significant part of the school curriculum is taught using French, the children's second language, as the language of instruction. French language instruction (the explicit teaching of vocabulary, grammatical rules, etc.) is not necessarily excluded from the immersion classroom, but it is the opportunity that immersion provides for children to learn language through its use rather than through explicit instruction which is often considered its distinguishing feature (Stern, 1978).

Literature Review

The most well-known type of immersion program is *early immersion*, which begins in kindergarten or grade one with children receiving their first two or three years of schooling with French as the only language of instruction. A considerable amount of research has been conducted on early immersion programs across Canada (Genesee, 1978; Lambert & Tucker, 1972; Shapson & Kaufman, 1978a; Swain, 1974, 1978a). This research has consistently demonstrated that children enrolled in such programs: (a) are not harmed in their cognitive or native language development; (b) master content subjects such as mathematics or science as well as their unilingually educated peers; and (c) develop high levels of proficiency in French.

More recently, *late immersion* programs starting at the upper grade levels have been implemented in Canadian schools. In Table 1 are listed some of the late immersion programs in Canada for which research information is available. As is evident, late immersion programs differ along several dimensions (e.g., grade of entry, amount of instructional time in French, student background in French). Despite the differences among the programs, research studies (Barik, Swain & Gaudino, 1976; Genesee, 1978/79; Shapson & Kaufman, 1978b) have been remarkably consistent in their findings which, to date, indicate that students enrolled in a late immersion program:

1. achieve at least as well as their peers in the regular English program in English, mathematics and other school subjects;[2]
2. perform at significantly higher levels in French than their peers enrolled in traditional French as a second language programs where French is taught as a separate subject for a specified number of minutes a week rather than used as the language of instruction; and
3. show positive attitudes toward French language and culture and toward their immersion experience.

More recent research on late immersion has investigated the relative merits of early and late immersion programs and has also yielded data on the ultimate (end of grade 12) levels attained by late immersion students. Although it is suggested in Montreal (Genesee, 1979) that a two-year late immersion program produces the same effects as an early immersion program in all areas of French language skills tested except French comprehension, research conducted in Ottawa/Carleton (Morrison, Walsh, Pawley & Bonyun, 1979) and in Peel County (Lapkin, Swain,

Kamin & Hanna, 1981) indicates that late immersion students do not equal their early immersion counterparts on most French language measures.

In order to determine the ultimate outcomes of late immersion programs, Morrison, Pawley and Bonyun (1979) investigated the French language skills of grade 12 students in Ottawa/Carleton who had been enrolled in different kinds of late immersion programs for from four to six years. The researchers found that in listening, reading, and writing, the average grade 12 student from these programs could reach the middle (Level B) of three levels of linguistic competence as defined by the Public Service Commission. However, the level attained by such a student in speaking was lower and was classified as Level A.

TABLE 1

EXAMPLES OF LATE IMMERSION PROGRAMS IN CANADA

Location	Grade of Entry	Student Background in French	Percentage of Instruction in French
Peel County, Ontario (Lapkin et al., 1981)	8	30 min/day in grade 6; 20 min/day in grade 7.	grade 8: 53%-70% grade 9: 40% grade 10: 40%
Ottawa, Ontario (Morrison et al., 1979a,b)	6	20 min/day from K-5	grade 6: 100% grade 7: 50% grade 8: 50% grade 9-12: 50%
Carleton, Ontario (Morrison et al., 1979a,b)	7	20 min/day from K-5	grade 7: 80% grade 8: 80% grade 9-12: 30%
Montreal, Quebec (Genesee & Morcos, 1978; Genesee, 1979)	7	15 min/day in K; 20-30 min/day in grades 1-6	Option 1: grade 7: 85% grade 8-11: up to 40% Option 2: grade 7/8: 85% grade 9-11: 20%
Coquitlam, British Columbia (Shapson & Kaufman, 1978b)	10	20 min/day in grades 6-7; 40 min/day (one semester) in grades 8-9	grade 10: 100% (One semester only)

Description of Late Immersion Programs in British Columbia

While a one-semester immersion program for secondary students has been offered in one school district in British Columbia since 1976 (Shapson & Kaufman, 1978b), programs for *elementary* school students have been introduced only very recently in the province. During the 1978/79 school year, late immersion programs were implemented in grade 6 by three school districts, all located in the Vancouver lower mainland, and they were continued to grade 7 in 1979/80. Although grade 6 was the entry point, the three programs differed in the amount and distribution of time in French and English, the subjects allocated to the two languages, the previous background of the students in French, and the screening of students for admission to the program (Table 2).

TABLE 2

LATE IMMERSION PROGRAMS IN BRITISH COLUMBIA

District	Percentage of Instruction in French		Student Background in French	Screening Procedures
	Grade 6	Grade 7		
#1	90%	60%	One year of EFSL (80 mins. per week in gr. 5)[a]	No
#2	60%	70%	One to two years of EFSL at various grade levels[a]	No
#3	60%	60%	None	Yes

[a] EFSL: Elementary French Second Language programs are traditional programs in which French is taught as a separate subject for a specified number of minutes per week.

In district 1, students were admitted to the program on a voluntary basis. Prior to entering the program, they had approximately 80 minutes per week of traditional elementary French as a second language (EFSL) instruction for one year in grade 5. In grade 6, French was the language of instruction for 90% of the curriculum for all subjects except physical education, music, and library, and English language arts was not offered. In grade 7, English language arts was reintroduced into the curriculum, and the amount of French instructional time reduced to 60%.

In district 2, students were also admitted to the program on a voluntary basis; most students entering the program had received one or two years of traditional EFSL instruction at some time during their elementary schooling. French was the language of instruction for approximately 60% of the curriculum in grade 6.[4] Students received English language arts, library, music, and band in English, while they received mathematics, science, social studies, French language arts, physical education, and art in French. In grade 7, the amount of French instructional time was 70%.

Unlike the students in districts 1 and 2, students in district 3 had no French prior to entering the program, and they had to meet the following criteria to be accepted for admission: (a) a rating by their teacher of average or above average in academic achievement; (b) performance at the grade 5 level or above on the English and mathematics sections of the Canadian Tests of Basic Skills; and (c) a genuine desire to participate in the program (student interviews). French was the language of instruction for 60% of the curriculum in grade 6 and grade 7. Subjects taught in French were mathematics, social studies, French language arts, music, and art. Subjects taught in English were English language arts, science, library, and physical education.

Summary of Previous Research on Late Immersion Programs in British Columbia

In the first year, evaluation studies were conducted of the three grade 6 programs in order to determine the effects of the program on student achievement in English and other school subjects, on their French language skills, and on

attitudes toward French language and culture (Shapson & Day, 1979a, 1979b; Pfeiffer, 1980).

Comparison groups consisted of children enrolled in the regular grade 6 classes in the same schools housing the late immersion classes and receiving approximately 60-80 minutes per week of elementary French second language instruction. Statistical analyses of the data were carried out to determine whether there were significant differences between the immersion and comparison groups on a variety of pre- and posttest measures such as cognitive abilities, French, English, and attitudes. Where appropriate, the pretest results were used to control statistically for initial differences between the groups.

The results of the studies are summarized as follows:

1. At the end of the year, the immersion students in all three districts performed at least as well as their peers in the regular program on tests of English vocabulary and usage, mathematics problem-solving, social studies, and science.
2. At the end of the year, the immersion students in all three districts attained a nearly perfect score on the B. C. French Comprehension Test, a test designed for students completing their second year of a traditional EFSL program. They also performed well on a more advanced test of French comprehension, the MLA French Comprehension Test, and their performance on this test compared favorably with that of grade 10 and 11 high school students completing four-week summer immersion French programs.
3. In two of the districts (district 1 and 3), the immersion students were found to have developed more positive attitudes toward French language and culture in comparison to their peers in the regular program.

Overall, these results were consistent with the pattern of results reported in research studies of late immersion programs elsewhere in Canada (e.g., Genesee, 1978/79; Swain, 1978a). Because students in the B.C. programs had little French instruction before their immersion experience and less previous French instruction than late immersion students elsewhere (Table 1), the results were viewed as being particularly encouraging. During the 1979/80 school year, all three districts continued the grade 6 program and provided a second year of immersion at grade 7, and the authors conducted a follow-up investigation of the grade 7 students' progress in French.

Description of the Present Study: A Comparison of the Three Programs

While research on late immersion programs has shown that students can acquire proficiency in French without apparent detriment to their native language development or achievement in other school subjects (Barik, Swain & Gaudino, 1976; Genesee, 1978/79), many questions still remain about the implementation of such programs. Among the questions are those related to the distribution of time in French and English, the subjects to be allocated to the two languages, the importance of previous instruction in French as preparation for the immersion experience, and the need for selection criteria. Because the three B.C. programs differed along these dimensions, the present study was undertaken to compare the three programs in order to gain comparative information about the student populations and to determine the relative effects of each program. The following questions were investigated:

Pretest (Beginning of Year 1):

1. How do the students in the three programs compare in their general cognitive abilities when they enter the immersion program (grade 6)?
2. How do they compare in their knowledge of French and attitudes toward French language and culture at the beginning of their immersion experience (grade 6)?

Posttest (End of Year 1)

1. At the end of the first year (grade 6), are the three programs equally effective with respect to student achievement in English vocabulary and usage, mathematics problem-solving, science, and social studies?
2. At the end of the first year (grade 6), are the three programs equally effective: (a) in developing positive attitudes toward French language and culture and (b) in developing students' French comprehension skills?

Follow-up (End of Year 2):

1. At the end of the second year (grade 7), are the three programs equally effective in further developing the students' French comprehension skills? Are they equally effective with respect to student achievement in French language arts (district 1 and 2 only)?

Methodology

Population

All the students who entered the late immersion program in the three school districts in grade 6 participated in the study. The three groups of students were as follows:

1. *District 1 Late Immersion Group*: The year 1 group was composed of 46 students enrolled in two late immersion classes receiving approximately 90% instructional time in French. The year 2 group was composed of 37 students in two classes receiving approximately 60% instructional time in French.[5]
2. *District 2 Late Immersion Group*: The year 1 group was composed of 28 students enrolled in one late immersion class receiving approximately 60% instructional time in French. The year 2 group was composed of 27 students (one class) receiving approximately 70% instructional time in French.
3. *District 3 Late Immersion Group*: The year 1 group was composed of 34 students enrolled in two late immersion classes receiving approximately 60% instructional time in French. The year 2 group consisted of 34 students (two classes) receiving approximately 60% instructional time in French.

Design and Analysis

The design of the study is illustrated in Table 3. In the first year of the study, a test of general cognitive abilities was administered near the beginning of the school year in order to have a baseline measure for comparing the groups. As well, at this time, students received a test of French designed for British Columbia students enrolled in elementary French second language programs and a Student Attitude Scale designed to measure their attitudes toward French language and culture.

The Student Attitude Scale was also administered at the end of the year in order to determine whether there were any changes as a result of the immersion experience. At the end of the year, students were given a more advanced test of

French comprehension (MLA French Comprehension Test) which was designed for high school students with one to two years of traditional French second language instruction. Tests measuring English language skills, mathematics problem-solving, science, and social studies were also administered toward the end of the year in order to measure achievement in school subjects.

In the follow-up study, students in all three districts were given the MLA French Comprehension Test toward the end of the year, and students in districts 1 and 2 also received a test of French language arts which was designed for francophone children in Montreal (Test de Rendement en Français).

Statistical analyses of the data were carried out to determine whether there were significant differences among the three groups at the beginning of their immersion experience in grade 6. In comparing the groups at the end of grade 6 and grade 7, an analysis of covariance procedure was used.

TABLE 3

DESIGN OF COMPARISON STUDY

Instruments	Grade 6 (Year 1)						Grade 7 (Year 2)		
	Pretest			Posttest			Follow-up		
	Dist. #1	Dist. #2	Dist. #3	Dist. #1	Dist. #2	Dist. #3	Dist. #1	Dist. #2	Dist. #3
Canadian Test of Cognitive Abilities	X	X	X						
Canadian Tests of Basic Skills									
English Vocabulary				X	X	X			
English Usage				X	X	X			
Mathematics Problem-solving				X	X	X			
Science Test				X	X				
Social Studies Test				X	X				
Student Attitude Scale	X	X	X	X	X	X			
B.C. French Comprehension Test									
Part A	X	X	X						
Part B	X	X	X						
Modern Language Association French Comprehension Test				X	X	X	X	X	X
Test de Rendement en Français (Level 1, Sec.)							X	X	

Measurement Instruments

A brief description of the measurement instruments used in the study follows:

The Canadian Cognitive Abilities Test, Level D (Thomas Nelson & Sons Ltd., 1977). This test is intended to measure the student's ability to use and manipulate abstract and symbolic relationships. Three separate batteries (verbal, quantitative, and nonverbal) are included to assess competence in working with symbols representing words, symbols representing quantities, and symbols representing spacial, geometric, or figural patterns.

Canadian Tests of Basic Skills, Form 3M Level 12 (Thomas Nelson & Sons Ltd., 1976). This multilevel battery of tests, standardized in Canada, is designed

to measure, in English, the development of skills in vocabulary, reading comprehension, work study, written language, mathematics concepts, and problem-solving. In order to minimize testing, only the following subtests of the CTBS battery were administered in the study: vocabulary, language usage, and mathematics problem-solving.

Social Studies Test. This instrument is made up of 12 items taken from the British Columbia Social Studies Assessment tests, which were developed and administered to grade four, eight, and twelve students in British Columbia public schools in April, 1977 as part of the Provincial Learning Assessment Program (Mussio & Greer, 1980). The 12 items were those common to the grade 4 and grade 8 tests of the B.C. Social Studies Assessment.

Science Test. This instrument consists of 28 items taken from the British Columbia Science Assessment tests, which were developed and administered to grade four, eight, and twelve students in British Columbia public schools in the spring of 1978 as part of the Provincial Learning Assessment Program (Mussio & Greer, 1980). As with the Social Studies Test, only those items common to grades 4 and 8 of the B.C. Science Assessment tests were included.

Student Attitude Scale. This is a 20-item scale assessing attitudes toward French language and culture. It was designed for use with students studying French in EFSL programs (Shapson, Kaufman & Durward, 1978). The reliability coefficients (Cronbach's Alpha) for the attitude scale are .91 (pretest) and .92 (posttest).

B.C. French Comprehension Test (Shapson, Kaufman & Durward, 1978). The B.C. French Comprehension Test was designed to measure the French comprehension skills of students in grades 4 to 7 who are completing their second year of a traditional EFSL program. The test, recorded on cassette tape, requires approximately 35 minutes to complete and is administered in one sitting. It consists of three parts: Sound Discrimination, 10 items; Vocabulary, 30 items; and Listening Comprehension, 18 items. The reliability coefficient (Cronbach's Alpha) is .91 for the test total.

Modern Language Association Cooperative Foreign Language Test, French, Form LA (Addison-Wesley Publishing Co., 1963). The MLA test is a standardized listening comprehension test designed for students in their first or second year of a high school French course. The test is presented on reel-to-reel tape and requires 25 minutes to complete. It contains six parts requiring responses to orally presented statements, questions, short conversations, radio broadcasts/ commercials, a telephone conversation, and an extended conversation.

Test de Rendement en Français (Level 1, Secondary). This 35-item test, developed by the Bureau de l'évaluation, la Commission des Ecoles Catholiques de Montréal, measures achievement in French language arts through subtests in spelling, vocabulary, reading comprehension, sentence completion, and grammar. Norms are available to make comparisons with francophone children of an equivalent grade level in Montreal.

Procedures

In all three districts, the pretesting was conducted in mid-October, and the posttesting in May of the school year. The pretesting (year 1) required a total of four sessions, with each of the three subtests of the Canadian Test of Cognitive Abilities administered in separate sittings and the B.C. French Comprehension

Test and Student Attitude Scale administered together in another sitting. The posttesting (year 1) also required a total of four sessions, as follows: (a) Canadian Tests of Basic Skills; (b) Science Test and Social Studies Test; (c) B.C. French Comprehension Test and Student Attitude Scale; and (d) MLA French Comprehension Test. In the follow-up, the MLA French Comprehension Test and the Test de Rendement en Français were administered in two separate sittings during the month of May.

In most cases, the testing was conducted by a member of the research team. However, in all three districts the classroom teacher administered the Canadian Tests of Basic Skills and in a few cases another instrument, according to standard procedures and instructions which were provided by the research team.

Results

Comparisons Among Groups on Pretest Measures (Year 1)

From the results of an analysis of variance comparing the mean scores of the three late immersion groups on the Canadian Cognitive Abilities Test, no significant differences were seen on any of the subtests or on the total test score (Table 4). This is an indication that the program in all three districts attracted groups of students of similar intellectual development.

A comparison of the performance of these students with the national norms for the Cognitive Abilities Test indicates that all groups performed at above average levels (69th percentile or above) on the verbal, quantitative, and nonverbal subtests. It is suggested that whether selection procedures were used in a district or not, students who opted for the late immersion program tended to be above average in intellectual development.[6]

While the students in the three districts did not differ in general cognitive abilities, in a comparison of their scores on the remaining pretest measures, significant differences were revealed in their level of French comprehension and attitudes toward French language and culture at the time of pretesting (see Table 4). Using post-hoc statistical comparisons at the .05 level, it may be seen that students in district 3 performed at a significantly lower level on the B.C. French Comprehension Test but had significantly more positive attitudes toward French language and culture than students in districts 1 and 2. These pretest findings appear to reflect the background of students in district 3 who started the program without any previous French instruction and who had been screened for their desire to enter the program.

It is noteworthy that after only 6 weeks in the immersion program, students in districts 1 and 2, who had a limited amount of previous French instruction, nearly reached the maximum level on the B.C. French Comprehension Test, Parts A and B (34.9/40 and 35.0/40, respectively). Students in district 3, who had no previous French instruction, attained a score which was comparable to that of students in the B.C. French Study (Shapson, Kaufman, & Durward, 1978) who had nearly three years of a traditional EFSL program (27.1 vs. 27.9, see Table 4).

At the time of pretesting, the attitudes toward French language and culture of all three groups of late immersion students tended to be more positive than those of students enrolled in traditional EFSL programs in the province, as can be seen in Table 4.

Comparisons Among Groups on Posttest Measures

The performance of the groups on the posttest measures of attitudes, achievement in school subjects, and French comprehension was analysed using the

analysis of covariance procedure (Table 5). The findings can be summarized as follows:

Achievement in school subjects. No significant differences were found among the groups on any of the subtests of the Canadian Tests of Basic Skills, such as English Vocabulary, Usage, Mathematics problem-solving, or on the Social Studies or Science tests.

French language skills. The results indicated no significant differences among the three districts on the MLA French Comprehension Test at the end of grade 6. Although students in district 3 had performed at a lower level on the B.C. French Comprehension Test (Parts A and B) at the beginning of grade 6, as suggested in the findings, they had overcome an initial lag in French.

Attitudes. While no significant differences were found on the French language measures, significant differences were found among the three groups in attitudes toward French language and culture. Indications from post-hoc statistical comparisons at the .05 level are that the students in district 2 had significantly less positive attitudes toward French language and culture than students in district 1 and district 3. While this may be an indication that the late immersion program in district 2 was less effective in improving students' attitudes toward French language and culture than were the programs in districts 1 and 3, it is important to note that the attitudes of the district 2 children remained in the *positive half* of the scale and that the lower attitude results were not paralleled by lower levels of performance on the tests measuring achievement in school subjects or French language skills.

TABLE 4

RESULTS OF PRETEST COMPARISONS (YEAR 1)

Pretest Measure	District 1			District 2			District 3			F-value	Comparative Data: B.C. French Study[a]	
	n	Mean Score	S.D.	*n*	Mean Score	S.D.	*n*	Mean Score	S.D.		*n*	Mean Score
Canadian Cognitive Abilities Test												
Verbal (max=100)	40	78.0	9.8	26	81.0	12.0	32	74.4	11.7	2.62 n.s.		
Quantitative (max=60)	39	45.0	6.3	28	44.0	7.3	32	43.9	5.9	0.33 n.s.		
Nonverbal (max=80)	39	69.7	8.0	28	71.1	6.1	32	70.7	5.7	0.34 n.s.		
Total (max=240)	37	192.9	19.2	26	196.2	21.9	32	188.9	19.9	0.96 n.s.		
B.C. French Comprehension Test												
(Total Part A & B, max=40)	41	34.9	3.4	28	35.0	2.7	32	27.1	3.6	64.07***	231	27.9
Attitudes toward French Language and Culture												
(20-item scale)[b]	26	44.8	8.4	28	42.8	11.2	32	36.6	6.6	6.84**	2824	48.1

n.s. Not significant
** Significant at the .01 level
*** Significant at the .001 level.

[a] The B.C. French Study (Shapson, Kaufman & Durward, 1978) was a study of traditional elementary French second language programs in British Columbia. The French Comprehension Test results are for a group of students who were tested toward the end of their third year of an elementary French second language program. The attitude results are pretest results for the grade 6 provincial group which participated in the B.C. French Study (first to fourth year of an elementary French program).

[b] Each item is rated on a five-point scale. The minimum or most positive possible score is 20 for the 20-item scale, the mid-point is 60, and the most negative is 100.

TABLE 5

RESULTS OF POST-TEST COMPARISONS (YEAR 1)

	District 1			District 2			District 3			F-value[a]
	n	Mean Score	S.D.	*n*	Mean Score	S.D.	*n*	Mean Score	S.D.	
Canadian Tests of Basic Skills										
Vocabulary (max=46)	41	31.5	6.6	26	34.6	7.4	32	30.9	7.4	0.93 n.s.
Usage (max=32)	41	22.9	4.1	26	23.8	4.7	32	21.6	3.9	1.68 n.s.
Math Problem-solving (max=31)	41	19.4	5.4	26	18.1	6.2	32	18.3	5.9	2.46 n.s.
Science Test (max=28)[b]	39	20.7	3.5	26	20.2	3.8	--	--	--	2.15 n.s.
Social Studies Test (max=12)[b]	40	10.0	1.1	26	9.8	1.3	--	--	--	0.49 n.s.
MLA French Comprehension Test (max=45)	41	25.4	6.9	26	28.6	6.6	32	27.5	6.5	21.60 n.s.
Attitudes toward French Language and Culture (20-item scale)[c]	22	41.7	9.5	27	49.6	13.5	32	39.0	8.7	6.61**

n.s. Not significant
** Significant at .01 level.

[a] Based on Covariate Analysis using the pre-attitude score as the covariate for the Attitude Scale and the Canadian Cognitive Abilities pretest score as the covariate for all other measures.

[b] Students in district 3 did not receive the Science or Social Studies Test.

[c] Each item is rated on a five-point scale. The minimum or most positive possible score is 20 for the 20 items, the mid-point is 60, and the most negative score is 100.

Follow-up Comparisons Among Groups on French Language Measures

The performance of the three groups on the MLA French Comprehension Test at the end of grade 7 was compared using the analysis of covariance procedure.

The three groups did not differ significantly in French comprehension at the end of grade 7 (Table 6). A similar analysis between children in district 1 and district 2 on the test of French language arts also revealed no significant differences on this aspect of French language skills. As was the case in the first year of the study, the programs were found to be equally effective in developing the French language skills of the students in the second year.

TABLE 6

FOLLOW-UP COMPARISONS ON FRENCH LANGUAGE SKILLS (YEAR 2)

	District 1			District 2			District 3			F-value[a]
	n	Mean Raw Score	S.D.	*n*	Mean Raw Score	S.D.	*n*	Mean Raw Score	S.D.	
MLA French Comprehension Test (max=45)	31	36.3	5.8	22	37.7	4.7	23	37.2	5.4	0.99 n.s.
Test de Rendement en Français Level 1, Secondary (max=35)	28	16.8	5.0	22	15.9	4.7	-	--	--	0.81 n.s.

n.s. Not significant

[a] Based on Covariate Analysis using the Canadian Cognitive Abilities pretest score as the covariate.

Summary and Discussion

Despite the variations among the three late immersion programs in student background in French, selection procedures, and amount of instructional time in French, the findings of this study are an indication that there were more similarities among them than there were differences, both in terms of the intellectual abilities of the students whom they attracted and in terms of student performance in the program. In the first year of the study, no differences were found among the students in the three programs in cognitive ability, in academic achievement, or in proficiency in French. However, at the end of the year, students in districts 1 and 3 were found to have developed significantly more positive attitudes toward French language and culture than students in district 2. In the second year of the study, students in all three districts again scored at similar levels in French comprehension and in French language arts (data for districts 1 and 2 only).

The similarity in achievement in French among the students in the three districts, particularly in the first year when the amount of instructional time differed substantially (90% for district 1 vs. 60% for districts 2 and 3), may seem surprising in view of previous research which has shown that there is a strong relationship between amount of instructional time and achievement in French. For example, this relationship has been demonstrated by Stern, Swain, McLean, Friedman, Harley and Lapkin (1976) in comparing different types of French programs, by Carroll (1975) in comparing the French language achievement of children enrolled in traditional French second language programs, and by Swain (1978b) in comparing early total and early partial immersion programs. In contrast, as in the present study, other researchers (Stern et al., 1976; Morrisson, Walsh et al., 1979; Lapkin et al., 1981) have also reported finding no differences in achievement in French among students enrolled in late immersion programs with different time variations. The students in the present comparison study tended to be above average intellectually, and it is possible that the relationship between instructional time and achievement in French is not as strong with this type of student as it appears to be when the range of student abilities is broader.

To date, much of the research on late immersion has been conducted on programs where there was a selection process (Barik & Swain, 1976) or where a process of self-selection operated (Bruck, Lambert & Tucker, 1976). While one should not conclude from this that students with a broader range of ability cannot profit from late immersion, it does indicate the need for caution in generalizing the results to a more intellectually diverse student population. Genesee (1978/79) has demonstrated that students of below average ability who participated in a grade 7 late immersion program in Montreal were not harmed in their academic development; they performed at similar levels on tests of English language arts and mathematics as did below average students in the regular English program. However, these students had received French second language instruction throughout elementary school prior to enrolling in late immersion.

The results of this study which involved direct statistical comparisons among three programs and which found few differences in student performance would seem to warrant the conclusion that a late immersion program can be equally effective under a variety of conditions. This is also supported by the stability across the variations of late immersion shown in previous research (Table 1). This may allow educators a certain amount of flexibility in adapting the programs to the needs and resources (e.g., administrative needs, availability of personnel) of

their particular school districts. However, further research is needed to determine whether this flexibility would be warranted in programs composed of students with a broader range of abiities.

Notes

1. It should be noted, however, that no direct statistical comparisons have been conducted among the studies.
2. Students in the grade 8 late immersion program in Peel Country, Ontario showed lags in science in the first year of the program when compared to the regular English program students; however, these initial lags disappeared by the end of the second year of the program (Barik, Swain, & Gaudino, 1976).
3. The Public Service Commission's *Language Knowledge Examination 400B Series* was administered to the late immersion students. This series is used to establish the A and B levels of linguistic competence. Another test is used to establish the highest of the three levels (Level C)); this test was not adminstered to the late immersion students.
4. Unlike the teachers in districts 1 and 3, who reported using English with the students only in exceptional circumstances, the grade 6 immersion teacher in district 2 reported that at the beginning of the year he used both French and English in teaching the students. The amount of French spoken was gradually increased and was estimated by the teacher to occupy 60% of instructional time by mid-year.
5. The 20% drop in enrolment in this program from year 1 to year 2 occurred largely because of family moves, changes of schools, etc. However, since the drop is greater than in Districts 2 and 3, it must be taken into account in considering the year 2 results.
6. Since these are comparisons between mean scores for *groups* and norms for *individual* students, the reader is cautioned that they can only be used as a general indication of level of performance.

References

Barik, H.C., & Swain, M. A Canadian experiment in bilingual education at the grade eight and nine levels: The Peel Study. *Foreign Language Annals*, 1976, *9*, 465-479.

Barik, H.C., Swain, M., & Gaudino, V.A. A Canadian experiment in bilingual schooling in the senior grades: The Peel Study through grade ten. *International Review of Applied Psychology*, 1976, *25*, 99-113.

Bruck, M., Lambert, W.E., & Tucker, G.R. Alternative forms of immersion for second language teaching. *Working Papers on Bilingualism*, 1976, *10*, 22-73.

Carroll, J.B. *The teaching of French as a foreign language in eight countries*. New York: John Wiley and Sons, 1975.

Cohen, A.D. The case for partial or total immersion education. In Simoes, A. (Ed.), *The bilingual child*. New York: Academic Press, 1976.

Genesee, F. A longitudinal evaluation of an early immersion school program. *Canadian Journal of Education*, 1978, *3*, 31-50.

Genesee, F. Scholastic effects of French immersion: An overview after ten years. *Interchange*, 1978/79, *9*, 20-29.

Genesee, F., & Morcos, C. A comparison of three alternative French immersion programs: Grades 8 and 9. Instructional Services Department. The Protestant School Board of Greater Montreal, December, 1978.

Genesee, F. A comparison of early and late immersion programs. Montreal: McGill University, 1979 (mimeo).

Lambert, W.E., & Tucker, G.R. *Bilingual education of children. The St. Lambert experiment*. Rowley, Mass.: Newbury House, 1972.

Lapkin, S., Swain, M., Kamin, J., & Hanna, G. Late immersion in perspective: The Peel Study. Toronto: The Ontario Institute for Studies in Education, January, 1981.

Morrison, F., Pawley, C., & Bonyun, R. French proficiency of graduates of high school bilingual programs, Ottawa and Carleton Boards, Spring 1979. Research Centre: The Ottawa Board of Education. French Working Paper No. 129, December, 1979.

Morrison, F., Walsh, M., Pawley, C., & Bonyun, R. French proficiency of Grade 8 primary- and late-entry immersion students, Ottawa and Carleton Boards, Spring 1979. Research Centre: The Ottawa Board of Education. French Working Paper No. 131, December, 1979.

Mussio, R., & Greer, N. The British Columbia assessment program: An overview. *Canadian Journal of Education*, 1980, *5*, 22-40.

Pfeiffer, M.G. *An evaluation study of the Grade VI bilingual programme in School District #35 (Langley)*. M.A. Thesis, Burnaby, B.C.: Simon Fraser University, January, 1980.

Shapson, S., & Day, E. *Evaluation study of the late immersion program: School District No. 43 (Coquitlam) Grade 6, 1978-79*. Burnaby, B.C.: Simon Fraser University, Faculty of Education, November, 1979. (a)

Shapson, S., & Day, E.M. *Evaluation study of the late immersion program: Surrey School District (Grade 6, 1978-79)*. Burnaby, B.C.: Simon Fraser University, Faculty of Education, November, 1979. (b)

Shapson, S., & Kaufman, D. Overview of elementary French programs in British Columbia: Issues and research. *The Canadian Modern Language Review*, 1978, *34*, 586-603. (a)

Shapson, S., & Kaufman, D. A study of a late immersion French program in secondary school. *The Canadian Modern Language Review*, 1978, *34*, 186-193. (b)

Shapson, S., Kaufman, D., & Durward, L. *B.C. French Study: An evaluation of elementary French programs in British Columbia*. Burnaby, B.C.: Simon Fraser University, Faculty of Education, July, 1978.

Stern, H.H. French immersion in Canada: Achievements and directions. *The Canadian Modern Language Review*, 1978, *34*, 836-854.

Stern, H.H., Swain, M., McLean, L.D., Friedman, R.J., Harley, B., & Lapkin, S. *Three approaches to teaching French*. Toronto: Ontario Institute for Studies in Education, 1976.

Swain, M. French immersion programs across Canada. *The Canadian Modern Language Review*, 1974, *31*, 117-129.

Swain, M. Bilingual education for the English-Canadian. In J. Alatis (Ed.), *Georgetown University round table on languages and linguistics, 1978*. Washington, D.C.: Georgetown University Press, 1978. (a)

Swain, M. French immersion: Early, late or partial? *The Canadian Modern Language Review*, 1978, *34*, 577-585. (b)

The Alberta Journal of Educational Research *Vol. XXVIII, No. 2, June, 1982, 149-162*

JUDITH CHAPMAN
Monash University

and

QUENTIN WILLIS
Royal Melbourne Institute of Technology

The Australian School Principal's Interpersonal World of Work

In the two studies reported here, it was indicated that interpersonal competence is fundamental for effective functioning in the principalship. Using the observational method to identify the work activity of principals, Willis found that the highest proportion of a principal's time was spent in direct, personal interaction with others. Using the survey technique to identify the competencies which principals need to possess, Chapman found that principals, teacher association representatives, school council presidents, and Education Department officials perceived personal qualities and interpersonal skills as the most important competencies required. Though conducted separately, both studies highlight the need for principals to possess the skills necessary to maintain healthy interpersonal relations in the work setting.

This paper is the product of two recent and independent studies of the Australian school principal. Willis (1980) used the direct observational method to describe the work activity of school principals and Chapman (1979) employed survey research to identify the competencies which school principals need to possess. The findings from these two studies highlight the prominence of interpersonal relationships in the work of the principal.

The concept of interpersonal competence and its relationship to organizational effectiveness was first developed by Argyris. Placing particular emphasis on the importance of interpersonal competence being possessed by people in leadership positions, he recommended:

Judith Chapman is a lecturer in educational administration at Monash University, Australia. Her background includes high school teaching, administration, and consultancy in organizational behaviour. Her research interests focus on principalship and educational planning.

Quentin Willis is a senior lecturer and coordinator of educational administration at Royal Melbourne Institute of Technology, Australia. His background includes high school teaching and administration and consultancy in organizational behaviour.

> More trust, concern for feelings, and internal commitment; more openness to, and experimenting with, new ideas and feelings in such a way that others could do the same. (Argyris, 1971, p. 18)

Problems in interpersonal relationships, Hunt (1972) noted, could be produced by conflicting expectations from organizational variables (such as the formal structure, the informal networks, and the technical sub-system) or from extreme pressures from the external environment.

Presently, in Australia[1], the school principal is embroiled in a variety of controversies over the educational process and, with augmented autonomy, is increasingly standing out as the nexus between the complex organization that is the school and the restless environment that is examining the school's relevance. At a time when the effectiveness of the school is under scrutiny, principals are experiencing the very pressures which place strain on their ability to interact competently with the individuals who constitute their school community.

Within the literature on organizational behaviour, there is substantial focus on interpretations and typologies of interpersonal relationships within organizational frameworks (Bennis, Schein, Berlew & Steel, 1964; Thompson, 1967; Whyte, 1969). However, scant literature is available specifically on the interpersonal relationships of school principals. Similarly, few data are available on the interpersonal competencies which principals need to possess.

In attempting to analyse the nature of the principal's work and in attempting to identify the competencies which principals need to function effectively, the two studies discussed separately in this paper produced findings which may help to develop a data base for clarifying the interpersonal characteristics of the principalship.

The First Method: Studying the Nature of the Principal's Work

Influenced by Mintzberg's (1973) study of five top American managers and O'Dempsey's (1976) observation of three Australian high school principals, Willis (1980) observed three school principals at work continuously for three weeks. They were the principal of a high school in the State system, an independent college conducted on behalf of a Protestant church, and a Catholic college run by a religious order in Melbourne, the capital of the State of Victoria, Australia.

So that the researcher could derive a comprehensive, first hand perspective of the principal's work in its fullest context, direct observation was chosen as the research procedure. A period of three weeks' observation for each principal was chosen. It was assumed that over the period of three weeks, the observer's influence would be diminished and the observer would soon merge into the background. Further, it was assumed that a minimum degree of participation by the observer would serve to restrict any control or manipulation of the principals' performance that might occur through interactive and evaluative feedback from the observer. The latter aimed to be only a "recording machine." There was a minimum of interaction between the observer and the subject being observed. What would be recorded would be public facts that could be observed and tested by anyone and constitute relatively impersonal knowledge (Van Dalen, 1973).

The three kinds of schools, from the main streams of secondary education in Australia — State High, Independent, and Catholic — would permit the drawing of a composite picture of the principals' work activities but at the same time would constitute a population from which generalising would be severely limited.

The basic criteria for the principals selected were that they be reasonably experienced in the job (at least five years as principal) and head schools of approximately 1000 students (to ensure a large, complex organization).

After the initial meetings with the principals, the observer arranged short "acceptance visits" as a pilot run for the main period of observation. It was made clear to each principal that prior approval had not been sought from the controlling or employing authority first. This was to ensure that the observer was not associated with the principals' "superiors" and that the study was not to be forced on them (Scott, 1965).

For 15 consecutive school days with each principal, the observer recorded data as the various work activities occurred. Record sheets were used: (1) to keep a chronological record of activities at work, (2) to catalogue incoming and outgoing mail, and (3) for the principals to keep a diary of after-hours work.

The emphasis of the findings was on the data from the observer's record rather than the more subjective diaries kept by the three principals. For each work unit or activity, the following variables were noted by the observer:

1. its duration;
2. the medium or means of performing the work;
3. its purpose;
4. its location;
5. the people involved;
6. who initiated the activity;
7. whether or not it was scheduled;
8. whether it was internal or external business;
9. whether it was interrupted and, if so, resumed.

The data for the total 2719 activities were reported in terms of two basic properties of the principals' work — the *duration* and the *frequency* of activities involved.

The observations were a mixture of structured and unstructured processes. Even though the activity form structured the data into categories such as duration and place, some items required only a mechanical response while others called for an element of judgment. However, each activity was described on the form essentially as it took place. There were 10 items on the activity form to be checked each time a separate activity occurred. In some cases there were 12, depending on whether or not the activity was interrupted and if it was the resumption of a previously interrupted activity. The values for each variable or work detail on the form were assigned in the post-coding stage after the observation period. Thus post-categorisation gave structure to the unstructured components of data collected. At first, it had been considered that categorisation could be done early in the observation period; however, not only did new categories continue to appear, but it was felt that such early structuring would limit the naturalness of the spontaneous description of the activities as they occurred.

Finally, the data were processed by computer, particularly to allow the gleaning of various cross-tabulations, such as the time spent by each principal with various other people, according to who initiated the activity and whether or not it was scheduled.

The study did not aim to make comparisons among the three principals. In fact, there was generally a high level of commonality in the findings of how they spent their work time. It is essentially the composite data that are reported here.

Findings

Over the total period of nine weeks' observation, the principals carried out a total of 2719 activities during 335.9 hours of work (averaging 37.3 hours per week) during the school day. Contact with other people occupied 75.4 percent of the principals' time at work. This aspect of the principals' work is shown in Table 1. Students, teaching, administrative staff, and executive staff in such positions as bursar, registrar, works manager, comprised those individuals internal to the school with whom the principal made contact. The principal also had plentiful contact with people external to the school who qualified the running of the school or who influenced its performance in some way: parents, former students, school council members.

TABLE 1

TIME SPENT ALONE AND WITH OTHER PEOPLE (INTERNAL & EXTERNAL) DURING THE SCHOOL DAY [a]

Category	Principals			Composite	
	A	B	C	Total	Average
Alone	29.3 (29.7%) 316 acts	27.1 (23.9%) 166 acts	26.4 (21.3%) 270 acts	82.7 752 acts	27.6 (24.6%) 251 acts
Others (Internal)	43.8 (44.2%) 584 acts	49.7 (43.8%) 365 acts	64.9 (52.7%) 614 acts	158.4 1563 acts	52.8 (47.2%) 521 acts
Others (External)	25.8 (26.0%) 189 acts	36.6 (32.5%) 94 acts	32.4 (26.2%) 121 acts	94.8 404 acts	31.6 (28.2%) 135 acts
TOTAL	98.8 (100.0%) 1089 acts	113.4 (100.0%) 625 acts	123.8 (100.0%) 1005 acts	335.9 2719 acts	112.0 (100.0%) 906 acts

[a] Hours over three-week periods.

The main medium of exchange used by principals was conversation. A total of 131.3 hours (39.1 percent) of their time at work was spent in conversation. These face-to-face conversations included activities such as a half-minute chat with students in the corridor, a five-minute talk with a deputy principal in order to relay information, and a 37-minute discussion with principals from feeder schools regarding enrolment figures for the following year. The composite average time spent on the media of work during the school day is featured in Figure 1.

The majority of contacts with people were made in the form of one-to-one interactions. The duration of contacts with only one other person added up to 124.6 hours (37.1 percent). Teachers were most frequently the other person in these personal encounters, with substantial time spent with the deputy principal (21.6 percent and 15.0 percent of the time with one person, respectively). A total

101.8 hours (30.3 percent) were passed with more than four people. Such activities, apart from committee or small group meetings, were mainly occasions when the principal was likely to be functioning in a ceremonial situation with hundreds of others or interacting briefly, as on a tour of the school. The composite average time spent by principals working alone and with numbers of other people is shown in Figure 2.

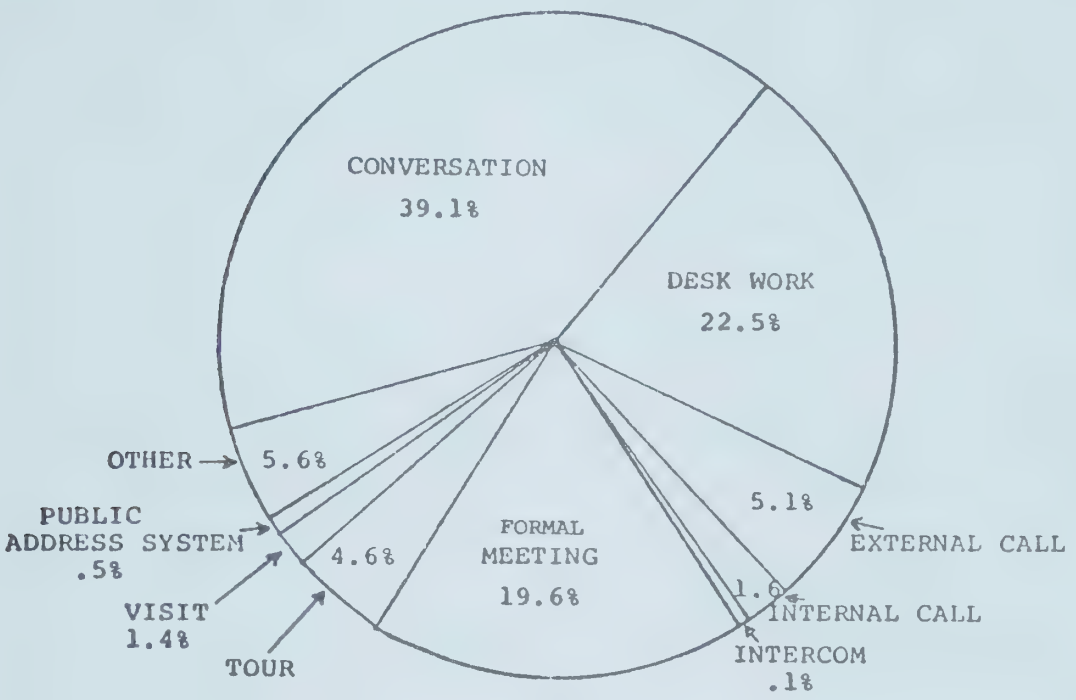

Figure 1. Composite average time spent on the media of work during the school day.

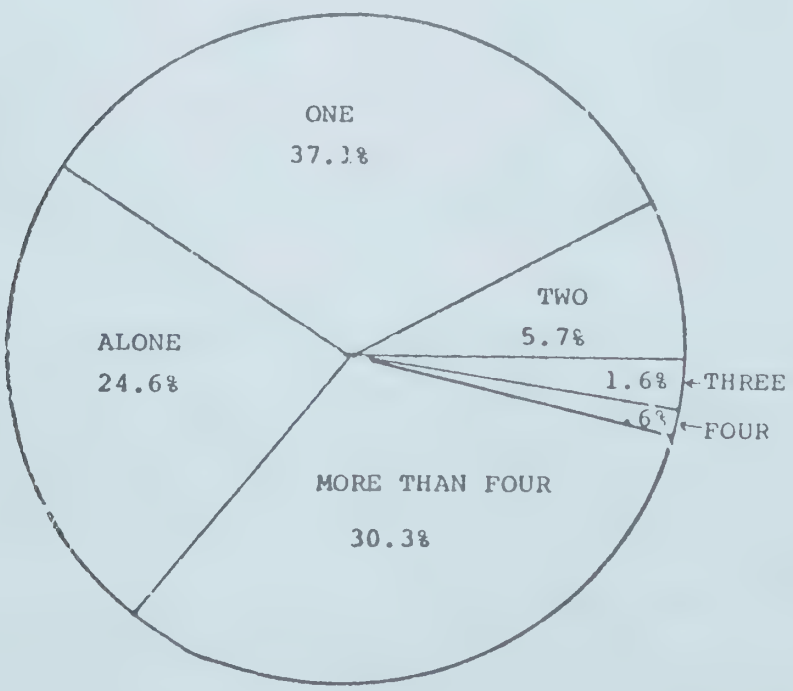

Figure 2. Composite average time spent working alone and with numbers of other people during the school day.

Information transmission was revealed as the major purpose of the activities of the principal. Altogether, 83.8 hours (24.9 percent) involved the principals in direct and obvious communication of knowledge. Thus, virtually one-quarter of principals' work during the day was in the role of giver and/or receiver of information about the life and business of the school. However, several other values for "purpose" also involved the passage of information, especially in terms of particular kinds of interpersonal contacts, such as planning, counselling, visibility, social conference, inspection, investigating, and even ceremonial. When added to the previous total for direct communication, these relationships show that at least a composite total of 218.3 hours (65.0 percent) was spent on some form of transmission of information. The purpose of the principals' work activities is indicated in Table 2.

TABLE 2

PURPOSE OF ACTIVITIES (COMPOSITE)

Category	Hours	Percentage
Informing	34.0	10.1%
Informed	42.8	12.7%
Information Exchange	6.9	2.1%
Planning	29.4	8.7%
Counselling	8.0	2.4%
Office Work	75.4	22.5%
Visibility	17.8	5.3%
Social	6.1	1.8%
Assembly	2.9	.9%
Conference	47.0	14.0%
Seek Information	2.0	.6%
Asked for Information	1.5	.4%
Inspection	13.5	4.0%
Trivia	3.8	1.1%
Ceremonial	9.2	2.8%
Teaching	8.5	2.5%
Investigating	3.6	1.1%
"Other"	23.7	7.0%
2719 Activities	335.9 hours	100.0%

In summary, it was in direct interaction with individual members of the school community that the principals spent most of their time. Such communication involved both the passage of information and interpersonal relationships. From the wide range of media available, the findings show that principals are most likely to be involved in the medium of face-to-face conversation. Skills in interpersonal relationships, Willis's findings imply, are necessary for effective functioning in principalship.

Though conducted quite independently from Willis', the research conducted by Chapman, in which constituents of the school system were asked to identify the competencies which school principals need to possess, provided further evidence of the importance of principals' possessing interpersonal skills.

The Second Method: Research in the Competencies Required to Perform the Tasks of School Principals

Influenced by studies on competency identification conducted by Cook and Van Otten (1972), Deros (1975), Gale (1973), McCleary (1975), Chapman (1979) used survey research to identify the competencies which school principals need to possess. Principals, Teacher Association representatives, School Council presidents, and Education Department officials (in Victoria, Australia) were asked to identify the competencies which they perceived as important to the principalship and those competencies which, though important, were perceived as not frequently displayed.

The competency statements which were listed in the survey instrument were generated using the informant technique. Initial contact was made with the Deputy Director General of Education. He informed the researcher of a wide range of educational personnel who could provide information relevant to the study. Contacts included the President and Vice-Presidents of the Principals' Association, principals of schools using traditional and innovative approaches, principals from schools of diverse socioeconomic surroundings, Education Department personnel, university faculty and teacher association representatives. Questions asked of the informants were designed to elicit information about the role of principals, their major tasks and duties, and the skills, knowledge, and personal characteristics perceived as necessary for principals to function effectively. Often-mentioned competencies were identified. From this source, the initial list of competencies was developed. This was reviewed in the light of literature containing competency statements identified as important to American secondary school principals. The initial list of competencies was refined and then given to a panel of administrators and academics to obtain face validity. Examples of ambiguity and repetition were corrected or culled; when agreement was reached on the meaning of a statement, that competency statement became a part of the final questionnaire.

The validated list contained 60 competency statements arranged in random order. In November 1978 a booklet containing the competency statements was mailed to the principal, Victorian Secondary Teacher Association (VSTA) representative, and School Council President in each of the 275 secondary schools in the State of Victoria. The list was also sent to senior officers in the Division of Secondary Education of the Victorian Education Department. The respondents were asked to rate each competency on two scales: a scale of importance (What Should Be) and a scale of performance (What Is).

On the scale of importance, respondents were asked to register their perceptions on the importance of each competency. The responses were registered on a four-point ordinal scale:

4 It is extremely important that principals possess this competency;
3 It is important that principals possess this competency;
2 It is not very important that principals possess this competency;
1 It is not at all important that principals possess this competency.

On the scale of performance, respondents were asked to register their perceptions of the frequency with which principals display the competencies. Again the responses were registered on a four-point ordinal scale:

4 Principals always effectively display this competency when it is appropriate;
3 Principals usually effectively display this competency when it is appropriate;

2 Principals rarely effectively display this competency when it is appropriate;
1 Principals never effectively display this competency even when required.

Responses were received from 118 principals (42 percent of principals throughout the state), 110 representatives of the VSTA (40% of VSTA representatives in the state), 81 School Council presidents, (29% of School Council presidents in the state), and 15 officers of the Education Department. Table 3 presents a comparison between respondents and the total population according to location of schools. Considering that in the State of Victoria, 54% of secondary schools are located in the inner city and suburban areas and 46% of secondary schools are located in country areas, respondents to the questionnaire in the principal and School Council categories proportionably represent the total population. Among VSTA representatives responding to the questionnaire, there was a slightly higher proportion of country representatives than would exist in the total population.

TABLE 3

COMPARISON OF RESPONDENTS TO TOTAL POPULATION ACCORDING TO LOCATION OF SCHOOL

	Percentage of Schools	Percentage of Respondents		
		Principals	VSTA Representatives	School Council Presidents
Inner City/Suburban	54.0%	51.0%	43.0%	52.0%
Country	46.0%	44.0%	53.0%	42.0%
Unspecified [a]	-	5.0%	4.0%	6.0%
TOTAL	100.0%	100.0%	100.0%	100.0%

[a] Percentage of respondents who did not specify the location of their schools on the questionnaire.

A frequency distribution was generated from the scores of each referent group for each of the 60 items. Intragroup agreement was defined as existing when 66% of respondents within a group agreed in their perceptions. Competencies considered important or extremely important by more than 66% of respondents in any of the four referent groups were also identified. The competencies were rank ordered according to their median scores on the scale of importance (What Should Be). A Spearman rank order correlation was used to determine the degree of relationship between the rankings of the referent groups. Scores on the scale of performance (What Is) were treated in the same manner. Responses on the scale of importance and scale of performance were compared. Competencies considered important and frequently displayed were identified as were competencies considered important but rarely/never displayed.

Perceptions of the Importance of Competencies

Forty-five competencies were perceived as important by respondents in all four referent groups. Personal qualities and interpersonal skills were the competencies

perceived to be most important by respondents in all groups. The competencies ranked most highly by respondents in all four referent groups are contained in Table 4.

TABLE 4

COMPETENCIES RANKED MOST HIGHLY BY RESPONDENTS IN ALL FOUR REFERENT GROUPS

	Scale of Importance: Median Score			
	Principals	VSTA Representatives	School Council President	Education Department Officials
Ability to establish rapport with school personnel through building feelings of mutual confidence and trust	3.87	3.87	3.78	3.80
Commitment to making the school a medium through which children can develop their potential	3.92	3.78	3.81	3.92
Ability to delegate authority to others	3.93	3.85	3.76	3.86
Ability to build a sense of cooperation and teamwork	3.80	3.77	3.78	3.96
Ability to make difficult decisions	3.87	3.75	3.80	4.00

These five competencies had been rated as extremely important by over 66 percent of respondents in all four referent groups. Skills such as interpersonal communication and group facilitation skills were also rated highly by respondents in all groups. All groups attached less importance to the skills involved in the operation and maintenance of the school's physical resources. Respondents' perceptions of the importance of skills necessary to maintain effective interpersonal relationships are contained in Table 5.

A significant relationship ($\alpha = .001$) existed between the rank orderings of each group. The strongest relationship existed between the perceptions of Education Department officials and VSTA representatives ($r = .992$) and between School Council Presidents and secondary school principals ($r = .992$).

Relationship Between Perceived Importance and Performance

A positive relationship existed among principals', Education Department officials', and School Council presidents' perceptions of importance and performance. No systematic relationship existed between VSTA representatives' perceptions of importance and performance.

TABLE 5

INTERPERSONAL COMPETENCIES CONSIDERED IMPORTANT BY RESPONDENTS IN ALL REFERENT GROUPS

	Scale of Importance: Median Score			
	Principals	VSTA Representatives	School Council President	Education Department Officials
Ability to maintain objectivity and fairness under pressure.	3.84	3.73	3.76	3.80
Ability to mediate situations of conflict among school constituents	3.83	3.75	3.79	3.72
Ability to maintain community confidence and support	3.72	3.39	3.62	3.75
Ability to project a positive image, generating enthusiasm and optimism in others	3.64	3.54	3.72	3.63
Ability to guide and direct meetings, maintaining participant interest and involvement	3.53	3.49	3.46	3.63
Ability in assisting others to formulate purpose and accept responsibility	3.40	3.24	3.46	3.38
Ability to express oneself effectively when addressing public gatherings	3.44	3.56	3.54	3.67
Skill in group decision making technique	3.36	3.43	3.26	3.56
Flexibility in dealing with people	3.35	3.43	3.32	3.20

TABLE 6

DEGREE OF RELATIONSHIP BETWEEN EACH GROUP'S PERCEPTIONS OF THE RELATIVE IMPORTANCE OF COMPETENCIES

	Principals	VSTA Representatives	School Council President	Education Department Officials
Principals		r=.886	r=.992	r=.882
VSTA Representatives			r=.926	r=.992
School Council Representatives				r=.916
Education Department Officials				

Six competencies were considered important and usually/always displayed by respondents in all four referent groups, as shown in Table 7.

TABLE 7

COMPETENCIES PERCEIVED IMPORTANT AND FREQUENTLY DISPLAYED BY RESPONDENTS IN ALL FOUR REFERENT GROUPS

	Frequency of Response: Usually/Always Displayed			
	Principals	VSTA Representatives	School Council Presidents	Education Department Officials
Ability to express oneself effectively when addressing public gatherings	94.1%	69.1%	81.5%	100.0%
Full acquaintance with the provisions of the Education Act and the rules and regulations of the Education Department	72.0%	80.9%	76.69%	66.7%
Ability to develop and maintain an efficient system of record keeping	82.2%	66.4%	77.8%	66.7%
Familiarity with the legal and economic limitations on decision making	81.4%	80.9%	66.7%	66.7%
Knowledge of the maintenance requirements of the school buildings and grounds	90.7%	78.2%	86.4%	86.7%
Knowledge of the organization and resources of the Public Works Department	80.5%	71.8%	75.3%	66.7%

These competencies have been traditionally associated with the principals' role as liaison between the school and the centralized offices of the Education Department. The personal qualities and interpersonal skills which were perceived as most important by all referent groups were perceived as displayed less often.

The widest discrepancy between importance and performance existed in the perceptions of teacher association representatives. Among the competencies which teacher association representatives perceived as important but not frequently displayed were two extremely important competencies: the ability to establish rapport with school personnel through building feelings of mutual confidence and trust, and the ability to build a sense of cooperation and teamwork. Teacher association representatives also perceived principals as rarely displaying the ability to generate enthusiasm in others, exert staff leadership, and facilitate group decision making. These competencies had been perceived as important by respondents in all four referent groups.

Conclusion

> The various functions which make up an organization are always mediated by the interactions of people, so that the organization can never escape its human processes. (Schein, 1969, p. 9)

The findings of Willis (1980) and Chapman (1979) indicate the importance of interpersonal competence in the work of school principals. Using the observational method to identify the work activity of school principals, Willis (1980) found that the highest proportion of a principal's time was spent in direct, personal interaction with other people. Similarly, when using the survey technique to identify the competencies which principals need to possess, Chapman found that principals, teacher association representatives, school council presidents, and Education Department officials perceived personal qualities and interpersonal skills as the most important competencies required of the principals. The findings support those of Wolcott (1973), Mintzberg (1973) and O'Dempsey (1976).

Though conducted separately, the studies of Willis and Chapman complement each other. Both reinforce the proposition that interpersonal competence is fundamental to effective functioning in the principalship. Both highlight the need for principals to possess the skills necessary to establish and maintain healthy interpersonal relations in the work setting. Both have implications for further research into the principalship — for example, to what extent does the health of a school's organizational climate depend on the principal's interpersonal competency?

Principals' preference for face-to-face conversation as the main medium for communication indicates the value principals attach to direct personal interaction. Willis' study indicated that many administrative tasks can be performed through interpersonal communication: through individual or small group meetings with teachers, the educational program can be monitored; through discussion with senior staff, the day to day running of the school can be supervised; through contact with community members, the school's relevance to the external environment can be maintained. The wide range of people with whom the principal must interact and the variety in types of interactions indicate the need for principals to possess an extensive repertoire of interpersonal competencies.

These interpersonal competencies, Chapman's study revealed, are among the most important that a principal can possess. While they are among those perceived as most important, they are not, Chapman found, among those perceived as most frequently displayed. This discrepancy between importance and performance indicates a need for principals to review their present practice in the field of interpersonal relationships. Though they may, as Willis has shown, be spending the majority of their time in direct personal interaction with others, the effectiveness of these interactions will certainly be reduced if interpersonal competencies are not possessed.

For those responsible for the preparation and selection of principals, the findings of Willis and Chapman have important implications. Chapman found that a high positive correlation existed among the perceptions of relative importance held by principals, school council presidents, teacher association representatives, and Education Department officials. This is an indication that groups are in agreement with regard to the competencies which they believe principals should

possess. The task of establishing priorities for selecting and preparing men and women for the principal position is thereby made easier. Priorities acceptable to one group will be acceptable to the others. Interpersonal competencies, such as those necessary to establish rapport, mediate conflict, and build co-operation, the results have shown, will be among those that have highest priority.

Notes

1. A clear trend in educational administration in Australia, especially in the large state systems, is the devolution of authority from the central administration to the schools and the consequent increase in the responsibility of the principal for running the school. At the same time, growing community involvement in education, as evidenced in the development of school councils, is directing pressure away from the central administration and onto the school principal.

References

Argyris, C. *Organization and innovation.* Homewood, Illinois: Irwin-Dorsey, 1965.

Argyris, C. *Management and organizational development.* New York: McGraw Hill, 1971.

Australian College of Education. *The community and education.* Papers of the Eighteenth Conference, Melbourne, 1977.

Baker, C. R. An observational study of a large public accounting firm. *Human Relations*, 1977, *30*, 11.

Barnes, R. M. *Motion and time study: Design and measurement of work.* New York: Wiley, 1963.

Bennis, W. G., Schein, E. H., Berlew, D., and Steel I. *Interpersonal dynamics: Essays and readings on human interaction.* Homewood, Illinois: The Dorsey Press, 1964.

Bickman, L. Observational methods. In C. Selltiz, L. S. Wrightsman & S. W. Cook (Eds). *Research methods in social relations.* 3rd ed. New York: Holt, Rinehart and Winston, 1976.

Boyd, R. D., & DeVault, M. V. The observation and recording of behaviour. *Review of Educational Research*, 1966, *36*, 5.

Chapman, J. D. *Competencies required in the performance of the tasks of Victorian secondary school principals.* Unpublished doctoral dissertation. University of North Colorado, 1979.

Cook, H. H., & Van Otten, K. *A study of the prime competencies required to perform the tasks of the secondary school principal.* Unpublished doctoral dissertation, University of Utah, 1972.

Copeman, G. How British executives spend their day. In G. Copeman, H. Luyk & F. de P. Hanika, *How the executive spends his time.* London: Business Publications, 1963.

Culbertson, J., Henson, Curtin & Morrison, P. (Eds.). *Performance objectives for school principals.* Berkeley, Calif.: McCutham Publishing Co., 1974.

Deros, L. *A study of the competencies required by Connecticut high school principals and those within the school system who influence his role.* Unpublished doctoral dissertation, University of Connecticut, 1975.

English, W., & Kaufman, R. *Needs assessment: A focus for curriculum development.* Washington, D.C.: Association for Supervision and Curriculum Development, 1975.

Festinger, L., & Katz, D. *Research methods in the behavioural sciences.* New York: Dryden Press 1953.

Gale, L. *Competencies required for the principalship.* Unpublished doctoral dissertation, University of Utah, 1973.

Glenn, J. R. *Chief executive time: An empirical study of the time allocation of American college and university presidents.* Unpublished doctoral dissertation, Stanford University, 1975.

Horne, J. H., & Lupton, T. The work activities of middle managers. An exploratory study. *The Journal of Management Studies*, 1965, *2*, 1.

Hunt, J. W. *The restless organization.* Sydney: Wiley, 1972.

McCleary, L. Assessing competency needs in administration. Paper presented at the meeting of the N.C.P.E.A. Interest Group on Competency Based Education for Administration, Bozeman, Montana, August, 1975.

Mintzberg, H. Structured observation as a method to study managerial work. *The Journal of Management Studies*, 1970, *7*, 1.

Mintzberg, H. Managerial work: Analysis from observation. *Management Science*, 1971, *18*, 2.

Mintzberg, H. *The nature of managerial work.* New York: Harper and Row, 1973.

O'Dempsey, K. Time analysis of activities, work patterns and roles of high school principals. *Administrators Bulletin*, University of Queensland, 1976, *7*, 8.

Peterson, K. D. The principals tasks. *Administrator's Notebook*, University of Chicago, 1978, *26*, 8.

Schein, E. H. *Process consultation: Its role in organisation development.* Reading, Mass.: Addison-Wesley, 1969.

Scott, W. R. Field methods in the study of organisations. In J. G. March, (Ed), *Handbook of organisations*, Chicago: Rand McNally, 1965.

Stellers, T. J. *Analysis of a suburban high school principal's time investment, activities and emphases.* Unpublished doctoral dissertation, Kent State University, 1973.

Thompson, J. D. *Organisations in action.* New York: McGraw-Hill, 1967.

Van Dalen, D. B. *Understanding educational research: An introduction.* New York: McGraw-Hill, 1973.

Whyte, W. F. *Organisational behaviour: Theory and application.* Homewood, Illinois: Irwin-Dorsey, 1969.

Willis, Q. F. *The work activity of school principals: An observational study in state, independent and catholic secondary schools in Victoria.* Unpublished doctoral thesis, The University of New England, 1980.

Wolcott, H. F. *The man in the principal's office.* New York: Holt, Rinehart and Winston, 1973.

Wolcott, H. F. The elementary school principal: Notes from a field study. In Spindler, G. D. (Ed.) *Education and cultural process: Towards an anthropology of education.* New York: Holt, Rinehart and Winston, 1974.

The Alberta Journal of Educational Research *Vol. XXVIII, No. 2, June, 1982, 163-174*

MONA BEEBE
and
GRACE MALICKY
The University of Alberta

An Exploratory Study into Determinants of Successful Reading Remediation

Not all children profit from remedial reading programs currently being provided. The results of an exploratory study in which an attempt was made to identify some of the determinants of successful reading remediation are reported in this paper. Specifically, the relationship between reading processes used by clients and the gains they made in silent reading comprehension were investigated. The sample consisted of 28 clients who had received remedial instruction in a University reading clinic. Unaided recalls were obtained and analysed for stories heard and for stories read silently both at entry to and exit from the program. The recalls were analysed to determine the nature of information recalled and the presence of logical connectives in the recalls. Results showed that students who were able to produce recalls that contained a high percentage of inferential information appeared to be more amenable to clinical remediation. In addition, use of perception connectives, which indicate understanding of time and space relationships, was related to gains in reading achievement.

Assessments of the effectiveness of remedial reading programs generally show significant short term gains. These results obscure the fact, however, that not *all* children are successful in these programs. To date little attention has been devoted to determining significant factors which differentiate those children who *do* succeed from those who *do not*. Such information would appear to be crucial for adapting remedial instruction to meet the needs of individual children.

Mona Beebe (PhD, University of Alberta) has taught in primary and elementary schools and, for one year, was a teacher of mother tongue English to primary and elementary school children in Sweden. Ms. Beebe has also taught courses in reading and language arts at the university level. Her research interests include the extension and elaboration of psycholinguistic models of reading.

Grace Malicky is currently a faculty member of the Department of Elementary Education at the University of Alberta. She teaches courses in reading education and is Director of the Reading and Language Center. She has been involved in clinical reading since 1968 and is currently conducting research in the area of adult literacy.

The overall purpose of this study, then, was to investigate determinants of success for children receiving remediation at the Reading and Language Center at the University of Alberta. Each year approximately ten elementary school students attend the Center where they receive diagnosis and remediation for their reading disabilities. Over the past two years, three-quarters of these children have benefited from the individual tutoring; some have gained as much as the equivalent of a two- or three-year reading level after only ten hours of remediation. One-quarter of the students, however, fail to benefit from such individualized treatment. A preliminary analysis of the data ruled out several factors as significant determinants of success; among these were age, grade, sex, intellectual functioning, educational level of parents, and number of schools attended. It was then hypothesized that the difference might be in *how* the children were reading when they entered the program, and attention was turned to an examination of reading processes utilized by a range of remedial students.

According to Drum and Lantaff (1977), good readers are able to remember the gist of what they read in either a free recall situation or when asked to recognize isolated components from the text. Memory for prose, they claim, is neither a verbatim translation nor a picture image of the input; rather, the gist that readers remember is the product of selecting and rearranging the elements within the text into a summary of the content of that text. The rearrangement and integration of the selected text information is believed to be based on the reader's prior knowledge, the anticipation of what the reader expects to find, and the way in which the text is structured as well as the reader's ability to interrelate items within and across sentences and paragraphs.

Tierney and Bridge (1979), among others, claim that by analysing a passage and comparing a subject's verbal recall of the text to the structure of the passage itself, the nature of the reader's organizing procedures can be assessed. Therefore, it was hypothesized that if remedial readers' recalls of passages read were analysed to try to determine *what kind* of information they selected from the text to rearrange and organize into a summary, as well as *how* they rearranged and organized this information, it might give an indication of what "gainers" are better able to do than "non-gainers" as they read.

In order to achieve this end it has been suggested by researchers such as Fagan (1978) that when considering the analysis of readers' recalls, an examination in at least two areas be conducted: (i) nature of the information recalled; and (ii) recall of logical connectives. The assumption is that the way the child subconsciously selects and organizes information as he *reads* the passage will be portrayed in the way he organizes and presents the gist of that passage in an unaided recall situation. A study of the kind of information that a child uses in his recall should shed light on what information he is selecting from the passage, how he is relating that information to prior knowledge in order to understand the passage, and how he is structuring the information (or misinformation) into a summary of the content of the passage.

Recall information is categorized by Drum and Lantaff (1977), Furniss (1979) and Fagan (1980) as falling into the following general categories:

1. information which is recalled verbatim from the text (text specific);
2. information which is inferred (text entailed);
3. erroneous and extraneous information (text elicited and/or text evoked); and
4. mazes, false starts, or repetitions (text external).

By analysing the pieces or units of information given in a reader's recall, one can estimate how much the reader is relying on trying to memorize exactly what is in the text and how much he is attempting to incorporate new incoming information into his already existing framework, a procedure which allows a reader to reproduce a recall that has a high degree of inferred or constructed information. Inferred information, then, may be thought of as requiring a higher order of processing or thinking than explicit or verbatim information.

The analysis of erroneous or extraneous information given by a reader during recall provides additional insight into processing activities. Cofer, Chmielewski and Brockway (1976) claim that "the knowledge that the reader possesses with which the passage he or she reads makes contact...is presumably the source of 'inaccuracy' that occurs during recall." Misinformation, then, may well be due to inadequate or inappropriate background for the passage being read.

Mazes, which include false starts, repetitions, corrections or voiced pauses (um, ah) appear to relate to a speaker's attempt to structure his output (Fagan, 1978, pp. 171-173). Rather than being viewed as errors or inappropriate utterances during a recall situation, mazes are now believed to possess various functions, some of which include planning of the next utterance, selecting a particular word and correcting or changing the wording.

According to the Fagan study, there also appears to be evidence from several studies (c.f. Drum, 1974; Frederiksen, 1975; Marshall, 1976) that information recalled from text is interrelated, rather than concatenated. Some of the linguistic devices which have been suggested as signalling these interrelationships are: referential terms (pronouns, synonyms, etc.); conjunctions (expressing a variety of relations); and staging or the positioning of information within a recall. Conjunctions convey the logical relationships that occur between informational units and are referred to by Fagan (1978, p. 97) as logical connectives. By examining the connectives the reader uses, an attempt can be made to observe how the reader is relating bits of incoming information to his existing background knowledge in order to provide a meaningful whole. Fagan (1980) has identified eleven different types of connectives which seem to fall into three groups:

1. descriptive or "what" connectives which depict relationships between equivalent units and the manner or mode of the relationship;
2. explanatory or "how and why" connectives which express relationships of cause and effect, reasons for events, agreements in favor of or against an event, contrastive events and conditions or limitations of events; and,
3. perceptual connectives which emphasize time and spatial dimensions of events.[1]

Purpose of the Study

This study was specifically designed to investigate the relationship between reading processes used by clients at entry into a remedial program and their gains in silent reading comprehension while in this program. Both the clients' ability to infer information during reading and their ability to use higher order connectives were considered to be indicators of processes used during their silent reading of narrative material.

In addition the study investigated the shifts in the use of these processes on pre- and posttest measures of reading comprehension for gain groups as compared to the group of no-gain clients.

Procedures of the Study

Sample

Subjects were selected from the Reading and Language Center files for the period 1976 to 1980. Only children in elementary and junior high schools (grades 2 to 9) with complete data on their pretest and posttest recalls were included in the sample. Twenty-eight subjects qualified for inclusion in the study. Six subjects made no gain or a gain of three months or less in silent reading comprehension after ten weeks of remediation. These students constituted group one. Group two contained eleven clients who had made a one-year gain. Group three also contained eleven clients who had progressed two or three years in their silent reading level. The mean reading achievement grade level score at the beginning of the remediation program for group one was 4.25, while for the second and third groups the mean grade level scores were 2.83 and 3.8 respectively.

Instrument

Several passages were read by each student during the administration of the *Standard Reading Inventory* (SRI) at both the diagnostic session and the final evaluative session. Two passage recalls for each of the 28 subjects were selected for analysis. Both recalls were unaided and were those obtained on passages that established the subjects' silent reading instructional level – one at the pretest or diagnostic session and the other at the posttest or final evaluative session.

At the time of the pretesting, the listening test from the SRI was also administered to most of the clients — 21 out of 28. These clients were asked to give unaided recalls of the listening passage in the same way that all clients were asked to give an unaided recall of the pre- and posttest reading passages. By administering the listening test, it was believed that an indication of the child's potential to organize and inter-relate pieces of incoming information could be obtained. Listening comprehension is generally considered by reading clinicians to provide one indication of reading potential.

Analysis of the Recalls

The first analysis conducted on the three recalls of each student was the identification of informational units. The procedures followed were those developed by Fagan (1978) for analysing children's oral language. The unit he suggested was a syntactic structure which may be of two types: a basic syntactic structure or an alternate structure. The basic structure is the simplest independent unit which may be used to convey information. For example, "Birds fly." or "Terry gave the dog a bone." are considered basic units. The alternate structure is appended to a basic unit and gives additional information. With the addition or substitution of words, alternate units could become basic structures. The interrater agreement between the assistant and the Director of the Reading and Language Center for dividing recalls into syntactic structures was .86 across four ratings when using the Arrington Reliability Formula (Fiefel & Lorge, 1950).

The second analysis consisted of allocating the informational units into one of four categories as adapted by Furniss (1979) from Drum and Lantaff (1977). The four categories, text specific, text entailed, text evoked, and text eternal, were discussed earlier in this paper. The interrater reliability score for the four main categories was .88, again across four ratings.

For the purposes of this study the first three categories constituted the total number of informational units. Text external units (mazes etc.) were considered as separate entities since they are basically tactics for mentally organizing

information that is about to be conveyed. Since passage recalls at the instructional level were based on different passages for most students, it was decided to convert the number of units each child gave in each category into proportions of the total recall units for that student. Hence, a child who gave five text specific units, ten text entailed units, and five text evoked units would have the following scores: text specific = .25, text entailed = .50, and text evoked = .25. In addition a ratio was established between the number of text external (mazes, etc.) and the total number of informational units. Thus if the child in the example repeated words, restated sentences, or added "ums" and "ahs" seven times, his ratio of external units would be 7:20 (or .35).

The final analysis consisted of categorizing the logical connectives between and within informational units into one of the eleven categories described by Fagan (1980, pp. 12-14). The connectives in the categories that made up each of the three larger groupings were then added together and again, proportional scores estimated such that each student had a proportional score for description connectives, for explanation connectives, and for perception connectives.

Limitations and Statistical Treatment of the Data

Several limitations in the case base made the use of inferential statistics in this study highly questionable. First, the size of the sample was small and may not have been representative of the reading disabled population in general. Second, different clinicians worked with each client since the Reading and Language Center is primarily concerned with training reading clinicians rather than with research. In addition, clients varied in age/grade levels as well as in instructional level, and hence, were administered different reading passages at different levels in the *Standard Reading Inventory.* This meant that the metric of the test instrument used for estimating a client's reading comprehension at one instructional level would be different than the metric of the test instrument used for estimating reading comprehension at another instructional level since the passages differed in length, difficulty and content. Finally, for 7 out of the 28 sample cases there was no information on listening; thus the sample was reduced insofar as this factor was concerned from 28 to 21 cases.

Because of the problems noted above, it was decided to limit the analysis to an examination of the emerging trends in the data by using means and zero-order relationships. It was felt that no single trend could be considered to constitute support for the hypothesis that children who entered the program with different language processing abilities benefited differentially from remedial instruction. Thus, the analytical strategy adopted was to assemble as many independent pieces of evidence as possible on the grounds that if the trends in these independent results were in the same general direction, then the collective evidence could hardly be ignored.

It is recognized that such a procedure is considered by many to be unorthodox in educational research since it is of a high inference variety. Nevertheless, it was necessitated by lack of prior information on the problem under investigation. Thus, the research was both exploratory and embryonic in nature and was designed to initiate inquiry into the problem as well as to stimulate further, more carefully controlled studies in the area.

Findings And Interpretation

Univariate Analysis

The first analysis of the results involved computation of group mean proportional scores to determine what trends emerged from the data.

Listening Comprehension. The results of the listening comprehension tests were analysed to provide some indication of the receptive language processes which readers possessed and which, therefore, had the potential to be used in the reading situation. The results of this analysis, shown in Tables 1 and 2, reveal the following trends.

1. The mean proportion of text specific recall units on the listening test was higher for the no gain and middle gain groups (.370 and .372 respectively) than for the high gain group (.281).
2. The mean proportion of text entailed or inferential units was .556 and .572 for the middle and high gain groups while for the no gain group it was only .383.
3. The no gain group gave more text evoked or erroneous and extraneous units (.247) than did the middle (.072) and the high (.147) gain groups.
4. Little difference existed on the productions of text external information units.
5. The no gain group used more descriptive connectives (.471) than the high gain group (.395) but fewer than the middle gain group (.505).
6. The no gain group used more explanation connectives (.226) than either the middle gain (.051) or high gain groups (.238).
7. The no gain group used fewer (.266) perception connectives than either the middle gain (.444) or the high gain (.343) groups.

Two trends appeared to emerge in the potential receptive language processes that the different gain groups brought to the remediation situation. First, the students that made gains during the 10-week remediation sessions were able, at the outset, to produce a higher proportion of inferential (text entailed) units in their recalls than were the students who made no gains. At the same time the no gain group gave more erroneous information as is evidenced by the higher proportion of text evoked units.

Second, the students who made gains used a higher proportion of perception connectives than did the no gain group. It would appear, then, that students who are the most likely to make gains after 10 weeks of remediation instruction are those who come to the remediation situation with the ability to perceive time and space relationships while listening to language. They are also able to relate what they hear to their own background knowledge and are therefore able to reproduce their understanding of the passage in inferential units.

Reading Comprehension. Clients were assessed at both entry into and exit from the remedial program. This made it possible: (1) to examine reading processes at two points in time; and (2) to examine changes (increases or decreases) in reading processes across time.

The following results were obtained on *pretests*:

1. During the pretest reading situation the no gain group gave fewer text specific (.260) and text entailed units (.380) than did either the middle gain group (.375 and .402 respectively) or the high gain group (.427 and .478).
2. No gain students gave more text evoked units (erroneous information) than did either group of gain students, .360 as compared to .223 (middle) and .095 (high).
3. High gain students tended to use fewer text external units (mazes, etc.) than did no gain or middle gain students.

TABLE 1

MEANS AND STANDARD DEVIATIONS OF LISTENING, PRETEST READING AND POSTTEST READING ON RECALL CATEGORIES BY GROUP MEMBERSHIP

Type of Recall	Recall Category	Group	Mean Number of Units	Mean Proportion of Total Units	S.D.	N
Listening	Text Specific (verbatim information)	Total Group	5.76	.332	.196	21
		No gain	7.60	.370	.218	5
		One year gain	6.71	.372	.232	7
		Two or more years gain	4.00	.281	.161	9
	Text Entailed (inferential information)	Total Group	9.57	.522	.210	21
		No gain	7.60	.383	.157	5
		One year gain	9.71	.556	.286	7
		Two or more years gain	10.56	.572	.145	9
	Text Evoked (erroneous information)	Total Group	2.71	.146	.134	21
		No gain	4.80	.247	.178	5
		One year gain	1.29	.072	.063	7
		Two or more years gain	2.67	.147	.123	9
	Text External (mazes, etc.)	Total Group	4.00	.220	.186	21
		No gain	4.60	.227	.130	5
		One year gain	3.14	.217	.151	7
		Two or more years gain	4.33	.218	.247	9
Pretest Reading	Text Specific (verbatim information)	Total Group	7.04	.371	.243	28
		No Gain	4.67	.260	.250	6
		One year gain	5.64	.375	.220	11
		Two or more years gain	9.73	.427	.262	11
	Text Entailed (inferential information)	Total Group	8.39	.427	.213	28
		No gain	8.50	.380	.221	6
		One year gain	4.82	.402	.220	11
		Two or more years gain	11.91	.478	.211	11
	Text Evoked (erroneous information)	Total Group	3.54	.202	.184	28
		No gain	6.33	.360	.155	6
		One year gain	3.45	.223	.208	11
		Two or more years gain	2.09	.095	.089	11
	Text External (mazes, etc.)	Total Group	4.96	.310	.398	26
		No gain	6.50	.363	.399	6
		One year gain	3.73	.374	.524	11
		Two or more years gain	5.36	.215	.127	11
Posttest Reading	Text Specific (verbatim information)	Total Group	5.00	.281	.146	28
		No gain	5.17	.296	.159	6
		One year gain	5.27	.322	.149	11
		Two or more years gain	4.64	.233	.137	11
	Text Entailed (inferential information)	Total Group	10.93	.545	.197	28
		No gain	12.17	.552	.191	6
		One year gain	7.18	.441	.166	11
		Two or more years gain	14.00	.647	.188	11
	Text Evoked (erroneous information)	Total Group	2.96	.175	.154	28
		No gain	2.33	.162	.175	6
		One year gain	3.45	.238	.165	11
		Two or more years gain	2.82	.120	.115	11

TABLE 2

MEANS AND STANDARD DEVIATIONS OF LISTENING, PRETEST READING AND POSTTEST READING ON CONNECTIVES BY GROUP MEMBERSHIP

Type of Connective	Connective Category	Group	Mean Number of Connectives	Mean Proportion of Total Connectives	S.D.	N
Listening	Descriptive	Total Group	7.43	.447	.162	20
		No gain	7.00	.471	.106	5
		One year gain	8.71	.505	.188	6
		Two or more years gain	6.67	.395	.169	9
	Explanation	Total Group	3.14	.189	.137	20
		No gain	4.00	.266	.111	5
		One year gain	1.00	.051	.058	6
		Two or more years gain	4.33	.238	.126	9
	Perception	Total Group	5.14	.353	.226	20
		No gain	4.40	.263	.129	5
		One year gain	7.29	.444	.199	6
		Two or more years gain	3.89	.343	.277	9
Pretest Reading	Descriptive	Total Group	7.04	.428	.219	28
		No gain	5.83	.382	.250	6
		One year gain	5.82	.429	.273	11
		Two or more years gain	8.91	.453	.145	11
	Explanation	Total Group	4.04	.240	.158	28
		No gain	4.83	.295	.199	6
		One year gain	3.09	.221	.172	11
		Two or more years gain	4.55	.229	.124	11
	Perception	Total Group	3.86	.252	.203	28
		No gain	2.83	.157	.182	6
		One year gain	3.09	.258	.279	11
		Two or more years gain	5.18	.297	.098	11
Posttest Reading	Descriptive	Total Group	7.71	.470	.207	28
		No gain	6.33	.463	.371	6
		One year gain	6.91	.440	.115	11
		Two or more years gain	9.27	.503	.177	11
	Explanation	Total Group	4.36	.299	.213	28
		No gain	5.50	.378	.337	6
		One year gain	3.27	.253	.122	11
		Two or more years gain	4.82	.303	.211	11
	Perception	Total Group	3.96	.269	.187	28
		No gain	4.00	.320	.351	6
		One year gain	4.64	.316	.115	11
		Two or more years gain	3.27	.196	.108	11

4. While the middle and high groups used more descriptive connectives (.429 and .453) than the no gain group (.382), the differences were not large.
5. Similarly, although the no gain group used more explanation connectives (.295) than the middle and high gain groups (.221 and .229) the differences were not great.
6. The greatest differences occurred in the use of perception connectives. The no gain group used fewer (.157) than either the middle gain (.258) or the high gain groups (.297).

When compared with listening comprehension, similar trends emerge. Students who were more amenable to clinical remediation were able to relate what they read to their background knowledge and thereby produce recalls that contained a high percentage (40 to 50%) of inferential information. Similarly, the ability to observe and use perception connectives during reading appeared to be of prime importance in predicting the likelihood of success during intensive individualized instruction.

An analysis of *posttest* results indicated the following:

1. The amount of text specific information given at posttest time decreased considerably from pretest time for the high gain group (from .472 to .233), less so for the middle gain group (from .375 to .322), but increased for the no gain group (.260 to .296).
2. The proportion of text entailed (inferential) units increased for all groups from pretest to posttest time, but the more substantial gains were for the no gain group (from .380 to .552) and the high gain group (.478 to .647). The high group, however, still gave considerably more inferential information at posttest time.
3. Concurrent with the increase in text entailed information for the no gain group was the decrease in the use of erroneous information (text evoked).
4. No great changes in the amount of text external information used was in evidence except, perhaps, by the middle gain group (from .374 to .246).
5. At posttest time all groups increased their use of descriptive connectives. The largest users were still the high gain group but the no gain subjects increased their use of these connectives the most (from .382 to .468).
6. Similarly, all groups increased their use of explanation connectives but only the no gain group made a substantial gain (from .295 to .378).
7. The no gain and middle groups increased their use of perception connectives at posttest time but the high gain group showed a decrease.

The trend of the importance of the students' ability to use inferential units of information during recall continued at posttest reading time. Conversely, an increase in verbatim recall in text information did not appear to be related to improvement in reading comprehension. In fact, the reverse was true. Those students who increased the use of verbatim information during the remedial sessions did not make gains in reading achievement, while those students who substantially decreased the use of verbatim information made the greatest gains. This may reflect greater attention to literal comprehension for the no gain group and a greater emphasis on inferential comprehension with those who did make gains. The posttest data were unclear in regard to use of perception connectives. Whereas this had been a good predictor of success at entry into the program, high gain subjects did not maintain their superiority in the use of these connectives.

Bivariate Relationships

Several of the findings based on the univariate statistics in Tables 1 and 2 were confirmed by the correlation analyses. For example, in Table 3, it is shown that the greater the use of inferential information in listening recalls, the greater the probability of group 3 (high gain) membership. Similarly, the greater the use of inferential information at the posttest level, the more likely it is that the student was from the high gain group. According to Table 3, the more erroneous (Text Evoked) information given by students at the pretest level, the more likely they were to be in the no gain group.

TABLE 3

ZERO-ORDER RELATIONSHIPS FOR RECALL CATEGORIES BY GROUP MEMBERSHIP

Variables		X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}
X_1	LSPEC	1.000												
X_2	LENT	-.783*	1.000											
X_3	LEVOC	-.231	-.425*	1.000										
X_4	LEXT	-.314*	.185	.168	1.000									
X_5	PRSPEC	.226	-.271	.094	.370*	1.000								
X_6	PRENT	-.503*	.556*	-.138	.457*	-.681*	1.000							
X_7	PREVOC	.338*	-.347*	.051	-.097	-.533*	-.257*	1.000						
X_8	PREXT	-.212	.129	.107	.225	-.080	-.007	.113	1.000					
X_9	POSPEC	.422*	-.342*	-.078	-.165	-.056	-.193	.296*	-.281*	1.000				
X_{10}	POENT	-.449*	.400*	.028	.184	-.095	.309*	-.232	-.126	-.628*	1.000			
X_{11}	POEVOC	.196	-.242	.094	-.070	.177	-.208	.007	.421*	-.149	-.674*	1.000		
X_{12}	POEXT	-.199	.147	.060	-.175	.143	-.192	.033	-.024	-.027	.014	-.000	1.000	
X_{13}	GROUP	-.205	.336*	-.226	-.016	.254*	.189	-.554*	-.163	-.203	.262*	-.164	-.082	1.000

* Significant at the $p < .10$ level.

Variable Identification: X_1=listening specific recall; X_2=listening entailed recall; X_3=listening evoked recall; X_4=listening external recall; X_5=pretest specific recall; X_6=pretest entailed recall; X_7=pretest evoked recall; X_8=pretest external recall; X_9=posttest specific recall; X_{10}=posttest entailed recall; X_{11}=posttest evoked recall; X_{12}=posttest external recall; X_{13}=group membership.

In terms of the relationships between connectives and gain scores, the most interesting finding was the importance of the use of perception connectives at the pretest reading level. It would appear that those students with the lowest competency in the understanding of time and space relationships were the least likely to achieve significant gains in reading comprehension due to intensive remediation.

Summary And Implications

Although no definitive conclusions can be drawn from this study because of limitations in the sample and in the test instruments, collectively the results support the thesis that the ability of problem readers to profit from remedial reading instruction involves integration of text information with background knowledge when both listening to and reading narrative material. This would seem to indicate that students who have been unable to profit from remedial instruction at the Reading and Language Center have generalized comprehension problems,

involving both reading and listening, rather than problems with specific reading processes per se. On the basis of these results it would appear that a broader, more integrated program involving both listening and reading comprehension is desirable for at least one subgroup of problem readers.

TABLE 4

ZERO-ORDER RELATIONSHIPS FOR CONNECTIVES BY GROUP MEMBERSHIP

Variables		X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}
X_1	LDESCR	1.000									
X_2	LEXPL	.020	1.000								
X_3	LPERCEP	-.708*	-.692*	1.000							
X_4	PRDESCR	-.174	-.154	.186	1.000						
X_5	PREXPL	.288	.234	-.364*	.076	1.000					
X_6	PRPERCEP	-.477*	-.123	.442*	-.359*	-.254*	1.000				
X_7	PODESCR	.240	.136	-.284	.191	-.056	-.413*	1.000			
X_8	POEXPL	-.141	.253	-.038	-.100	.097	.179	-.783*	1.000		
X_9	POPERCEP	-.182	-.157	.256	-.159	-.254*	.368*	-.572*	.356*	1.000	
X_{10}	GROUP	-.230	.026	.100	.121	-.139	.253*	.086	-.098	-.287*	1.000

* Significant at the $p<.10$ level.

Variable Identification: X_1=listening description connectives; X_2=listening explanation connectives; X_3=listening perception connectives; X_4=pretest description connectives; X_5=pretest explanation connectives; X_6=pretest perception connectives; X_7=posttest description connectives; X_8=posttest explanation connectives; X_9= posttest perception connectives; X_{10}=group membership.

The results of this study also support a stronger emphasis on inferential as opposed to factual comprehension with problem readers. Verbatim recall of information in stories heard or read was less related to improvement in silent reading achievement than was ability to relate background knowledge to text information. Results generally confirmed, as Tierney, Bridge and Cera (1979) have noted, that reading is both an abstractive and constructive process. However, although this was true for both high and no gain groups, a major differentiating factor involved the ability of high gain subjects to construct meaning which was consistent with the text information. Hence, more inferential information appeared in the recalls of high gain subjects than in those of the no gainers who generally produced more erroneous information. It should be noted that there was some shift away from erroneous inferential information in the recalls of the no gain group, suggesting that long term instruction focussed on inferential comprehension may have resulted in this group making achievement gains as well.

The role of connectives as a determinant of success in the remedial program examined in this study was less clear-cut. Ability to perceive time and space relationships at entry into the program appeared to be related to children's ability to profit from the instruction provided. However, posttest results did not indicate that heavy use of perception connectives was maintained by the high gain group. Again it appeared that the no gain group showed improvement in this area but that this was not yet reflected in gains in reading achievement. Until the relationship of connectives to achievement is clarified, however, special attention

to time and space relationships would appear to be merited in remedial programs for all problem readers.

Because of the descriptive nature of this study it has generated more hypotheses than conclusions. Studies are needed to ascertain whether or not the determinants of success identified in this study have predictive validity with a larger, more representative sample of problem readers and with informational as well as narrative materials. Research is also needed to assess the appropriateness of the program implications drawn from the results. What the study clearly supports, however, is the need for attention to be directed toward process differences (differences in *how* readers interact with print) in both research and practical work with problem readers. This study has identified only two of many potential process differences between subgroups of problem readers. The ultimate goal of continued work in this area is greater differentiation of instruction so that all problem readers will profit from remedial instruction.

Notes

1. Readers interested in a more detailed description of the procedures for identifying informational units, for categorization of the units, and for identification and categorization of logical connectives can obtain this information from Dr. W. T. Fagan, Department of Elementary Education, University of Alberta.

References

Cofer, C.N., Chmielewski, D.L., & Brockway, J.F. Constructive processes and the structure of human memory. In C.N. Cofer (Ed.), *The structure of human memory*. San Francisco: W.H. Freeman, 1976.

Drum, P.A. *Between sentence factors and the effects upon reading rate and recall.* Unpublished doctoral dissertation, Stanford University, 1974.

Drum, P.A., & Lantaff, R.E. Scoring categories for protocols. Paper presented at the Second Annual Language Conference, Boston University, October, 1977.

Fagan, W.T. *A comparison of the oral language of children, ages nine, ten and eleven.* A Research Report, Canada Council Grant, University of Alberta, 1978.

Fagan, W. T. Logical connectives. Unpublished paper, University of Alberta, 1979.

Fagan, W. T. Comprehension categories for protocal analysis. Unpublished paper, University of Alberta, 1980.

Fiefel, H., & Lorge, I. Qualitative differences in the vocabulary responses of children. *Journal of Educational Psychology*, 1950, *41*, 1-18.

Frederiksen, C.H. Effects of context-induced processing operations on semantic information acquired from discourse. *Cognitive Psychology*, 1975, *7*, 139-166.

Furniss, E.R. *Schemas for reading and recall of story narrative and descriptive informational texts: A study of sixth grade proficient readers.* Ph.D. dissertation, University of Alberta, 1979.

Marshall, N. The structure of semantic memory for text. Unpublished doctoral dissertation, Cornell University, 1976.

Tierney, R.J., Bridge, C., & Cera, M.J. The discourse processing operations of children. *Reading Research Quarterly*. 1979, *14*(4), 539-573.

The Alberta Journal of Educational Research *Vol. XXVIII, No. 2, June, 1982, 175-187*

VERNER R. NYBERG
and
S. C. T. CLARKE
The University of Alberta

School Subjects Attitude Scales

The purpose of the project was to develop and validate a semantic differential type instrument for measuring students' attitudes toward school subjects. At the completion of five developmental stages, the instrument comprised 24 bipolar word pairs with three factors or dimensions: evaluation, usefulness, and difficulty. Test-retest reliabilities on the various dimensions for various subjects and grade levels ranged from .61 to .86. Validity studies at the Grades 5 and 8 levels showed that scores on the three dimensions were predictably related to teachers' opinions, students' categorizations of school subjects, type of school, cultural differences, and sex differences.

Recent changes in school curricula have included emphasis on goals which involve the affective domain. The implementation of such goals, and of the teaching methods designed to promote them, could be furthered by instruments for measuring students' attitudes toward school subjects. Such considerations sparked the Alberta Minister's Advisory Committee on Student Achievement to recommend that an instrument be developed for this purpose. This article is based on the detailed report of its development (Nyberg & Clarke, 1979).

Ideally, the instrument should include an optimal admixture of desirable, but to some extent, conflicting features. It should be: (a) readily understood by teachers and students (that is, have face validity), yet not subject to faking; (b) valid and reliable, yet short enough so that several school subjects could be rated in a normal classroom period; (c) if possible, suitable for all school subjects at widely different grade levels; and (d) handscorable.

It seemed likely that an optimum blend of these desired qualities could be obtained by using the semantic differential technique. This inclination was

---

In 1966, Dr. Nyberg left a position in testing and research with the Alberta Department of Education to join the staff of the Department of Educational Psychology at the University of Alberta. His field work, many studies, and publications have been largely in the area of measurement.

Dr. Clarke still teaches and works in the Department of Educational Psychology following his retirement in 1977. The two men have recently collaborated in a number of large scale studies.

strongly supported by the need, initially, for structuring the attitude domain with respect to school subjects. No set of preconceived notions about this structure was required. Rather, the study would seek to delineate one through an analysis of the data.

The article by Maguire (1973) served as a guide for the rationale and methodology of this study. In the description of procedures used in this study, it will be seen that care was taken in selecting the scales; a variety of concepts (school subjects) was used, but no students were given an overwhelming task of differentiation; subjects of varying backgrounds were chosen, but no definition of a population for sampling purposes was attempted at this point; a "stringing out" procedure (Maguire, 1973, p. 301-2) was used for converting the data into a matrix; and a factor analysis method with varimax rotations was employed.

Procedures Used

Osgood, Suci, and Tannenbaum (1957) originated the semantic differential technique to study the meaning of concepts. Factor analyses of the correlations among responses usually yield three dimensions: *evaluation* (associated with bipolar adjective pairs such as pleasant-unpleasant), *potency* (associated with pairs such as strong-weak), and *activity* (associated with pairs such as fast-slow). It was anticipated that school-aged children and youths, rating school subjects in the affective area, might also produce results with these or similar factors.

The semantic differential technique has been widely used and has been applied beyond the study of concepts. For the present study, it was decided that a five-point scale would be suitable for children and youth, and that the ratings requested should be "very much, a bit, neither, a bit, very much." The scoring adopted was +2, +1, 0, -1, -2 respectively, with high positive scores signifying a positive attitude. The bipolar pairs of words used should be applicable to school subjects.

The general procedure involved stages of instrument development and modification followed by field trials, with each stage building on the results of the preceding one. Five stages eventuated.

Stage 1

This stage commenced with a search of the literature for bipolar pairs of words, usually adjectives, which might be applied to school subjects. Approximately 100 pairs were found; these were reduced to 53 by eliminating pairs such as warm-cool which did not appear to apply to school subjects. These 53 were then administered on a one-to-one basis to a few students, who were encouraged to talk out loud as they attempted to rate a school subject using either pair. This procedure eliminated seven pairs which were either not deemed at all applicable to school subjects or were not comprehended by the students. The remaining 46 pairs were arbitrarily assigned to two instruments of 23 pairs each.

Stage 2

At this stage there were two major purposes: to reduce the number of bipolar pairs of words, and to ascertain what factors were present in the students' responses. Maximum variability of response would further both purposes. Each instrument was administered to 441 Grade 5 students in 12 classes from 8 schools, and 307 Grade 8 students in 12 classes from 5 schools. Three teachers were chosen

and especially engaged for the task of administering the instruments. Each classroom group rated two subjects chosen by the administrator from the standard list (English language, English literature, reading, social studies, science, mathematics, health) or from a longer list of optional subjects such as art, music, and industrial arts. In any given school, only a few options were offered.

Each of the ratings, regardless of grade level and subject rated, was treated as input data for a factor analysis. Three factors emerged after rotation to a varimax solution. The first, labelled *evaluation*, was the strongest, with eleven pairs in the following descending order of loading: bright-dull, exciting-tiresome, lively-listless, like-dislike, pleasant-unpleasant, (nice-awful & dead-alive, tied), (happy-sad, alert-dull & boring-interesting, tied), still-active. The next, labelled *usefulness* included the following seven pairs: important-unimportant, valuable-worthless, necessary-unnecessary, useful-useless, meaningful-meaningless, smart-stupid, bad-good. The third, labelled *difficulty*, included six: easy-hard, hard-soft, clear-confusing, light-heavy, familiar-strange, certain-uncertain.

The three factors named above accounted for 57 percent of the variance on each of the two instruments, and were used in the next stage.

Stage 3

The purposes at this stage were two-fold: to ascertain whether the factors tentatively identified in the preceding stage were stable, and to explore the possibility of using a shorter instrument consisting of only twelve bipolar pairs of words.

The instrument which resulted from Stage 2 was administered by two of the three teachers who administered the first scale. There were 437 students in 12 Grade 5 classrooms from 9 schools and 307 students from 12 Grade 8 classrooms in 6 schools, each of whom rated social studies and one other subject chosen by the administrator. The set of schools and classrooms involved in this stage overlapped but were not identical with the set involved in Stage 2.

In all, 1474 ratings were useable. The three factors identified in Stage 2 remained. Based on the experience so far, minor revisions in the instructions to students were made. For each factor, the order of the bipolar pairs was altered to ensure that the positive end appeared at the right and left of the five-point scale an equal number of times. For each factor, the four word pairs with the highest loading were considered for the 12-pair short form. However, it was decided that the 24-pair longer form should be used for the next stage.

Stage 4

The results so far being sufficiently promising, it was decided to explore reliability and validity. For this purpose, one of the authors (Clarke) administered all of the instruments. Schools with certain known characteristics were carefully chosen. One school, in a large system, was an *academic* school. Parents who desired a distinctly academic emphasis could bring their children from all parts of a large city to this school. Another two schools were chosen because approximately one-third of the students were from a different cultural background. The remaining schools were chosen because proximity would facilitate repeated measures.

In each instance, teachers were asked to identify approximately one-quarter of the students who had the most positive attitudes toward social studies, and

approximately one-quarter with the least positive attitudes. About one-third of the students were asked to identify and rate the subject liked most, then the one liked least. Others were similarly asked to identify and rate the most useful and the least useful, while others were asked to do the same for the most difficult and least difficult. Where available, achievement and intelligence test scores were collected.

In two of the schools, one month later, the same students again rated the subjects which were originally rated.

Stage 4 involved 206 Grade 5 students in 9 classrooms from 3 schools, and 351 Grade 8 students in 15 classrooms from 3 schools. None of these schools had been previously involved in Stages 2 and 3.

Reliabilities were calculated for each of the factors: evaluation, usefulness, and difficulty, using results from all subjects rated. In the *evaluation* factor, coefficient alpha at the Grade 5 level based on 823 useable ratings was .956 for the long form (24 bipolar pairs) and .914 for the 12-pair shorter form; for *usefulness* the corresponding figures were .910 and .894, and for *difficulty*, .823 and .801. The corresponding figures for the Grade 8 level based on 1397 useable returns were: for *evaluation*, .961 and .909; for *usefulness*, .914 and .905; and for *difficulty*, .831 and .814. As expected, the shorter form was less reliable.

Test-retest reliabilities were calculated for those students who rated a subject twice at one sitting because it was assigned by the administrator and subsequently was chosen by the student as most-liked/least-liked (or most-useful/least-useful, or most difficult/least difficult). The number of students involved was low in many instances. The clearest results were for science at the Grade 8 level, where 107 students provided a test-retest reliability on the short form of .88 for *evaluation*, .90 for *usefulness*, and .80 for *difficulty*.

Test-retest reliabilities after one month were considerably lower than the figures illustrated above, ranging from a low of .37 on *difficulty* for 89 Grade 5 students rating the subject science (12-pair short form) to a high of .68 on *evaluation* for 119 Grade 8 students rating the subject science (24-pair long form, also 12-pair short form). Fluctuations of attitudes were suspected.

Because the scale was further revised, and results of further calculations of reliability are shown later, additional detail is not provided here.

Although the short form was only marginally less reliable than the long form, it was noted that the long form took only ten or twelve minutes to administer and about half of this time was devoted to handing out forms and to giving instructions. It was therefore decided to abandon the development of a short form.

The validity of the scales was assessed by obtaining evidence from five areas as described below.

Expert opinion was obtained by asking the teacher to identify about one-quarter of the students in a classroom who had the most positive, and another one-quarter who had the least positive attitudes toward social studies. At the Grade 5 level, 52 students identified as having a positive attitude had the following mean scores: *evaluation*, 2.73; *usefulness*, 4.67; and *difficulty*, 1.06. The corresponding scores for the 47 identified as having least positive attitudes were -0.06, 1.17, and -0.81. The differences between the means were all in the expected direction. The differences for *evaluation* and *usefulness* were significant at the $p<.01$ level, and for *difficulty* at the at the $p<.05$ level. For corresponding groups (86 and 76) of Grade 8 students, the differences between the corresponding means

were also all in the expected direction and, as before, the first two were significant at the $p<.01$ level and the last one at the $p<.05$ level.

Student ratings agreed with teacher opinion.

Student preferences provided a second basis for comparing mean scores on each factor. In about one-third of the classrooms, each student was asked to select the most liked subject and to rate it, then to select and rate the least liked subject. In other classrooms a similar rating was made of subjects deemed most useful and least useful, and most difficult and least difficult. The wide range of choices available made the numbers very small, so Grade 5 and 8 results were combined regardless of school subject. The mean score on the *evaluation* dimension for 250 Grade 5 and 8 students who selected various school subjects as most liked was 6.44, on *usefulness*, 5.35, and on *difficulty* 3.24. The corresponding scores made by the same students on school subjects selected as least liked were -4.89, -0.17, and -2.65. The largest difference was for the evaluation dimension. Each of the differences was significant at the $p<.01$ level. Similar results were found for the usefulness dimension, where the largest difference between mean scores was for that dimension, and each of the differences was significant at the $p<.01$ level. Similar results were found for the difficulty dimension.

It was clear that students' choices, or statements about subjects, coincided with their ratings of the subjects.

Analysis by type of school provided mean scores from an "academic" school. This school had only one Grade 5 class, composed almost wholly of children from the neighborhood who were not necessarily "academic." However, the three Grade 8 classes were, to a large extent, composed of students conveyed by their parents in order to attend this particular school. Comparing mean scores of students attending this school with those attending the remaining schools would provide a basis for validation. The comparison involved different schools with a small number of teachers who might conceivably use widely different methods of teaching. However, the Grade 8 students in the "academic" school were a select group. Statistics for the Grade 5 group were not computed because of a problem in administration that resulted in a very few students completing science or social studies ratings. Results for the Grade 8 classes are shown in Table 1.

TABLE 1

COMPARISON OF MEAN RATINGS OF GRADE 8 STUDENTS IN AN ACADEMIC SCHOOL WITH THOSE IN OTHER SCHOOLS

School Subject	Dimension	Academic Students			Other Students			Difference Between Means	Level of Sig. of Diff.
		N	$\overline{X}_1$	S	N	$\overline{X}_2$	S	$(\overline{X}_1-\overline{X}_2)$	
Science	Evaluation	61	-.33	5.51	284	.67	4.84	-1.00	-
Science	Usefulness	61	3.02	5.10	284	3.11	4.50	- .09	-
Science	Difficulty	61	-1.22	3.93	284	-1.46	3.86	.24	-
Social St.	Evaluation	62	2.76	3.27	285	.23	3.87	2.53	< .01
Social St.	Usefulness	62	4.71	3.62	285	3.41	4.15	1.30	< .01
Social St.	Difficulty	62	.31	2.76	285	-.57	3.37	.89	< .05

It is quite clear from the figures in Table 1 that the selected academic students rated science about the same as unselected students, but that for social studies the academic students recorded more positive attitudes in all three dimensions. These results were difficult to interpret because of confounding variables, such as the fact that the science students were exposed to a different, more rigorous science curriculum. In any case, it was not felt that the results cast any real doubt on the validity of the instrument.

Cultural differences were used as another criterion for studying validity. Students of a distinct subculture were regarded by teachers and administrators as having attitudes toward school subjects which were less favourable than those of students from the main culture. The comparative scores of these groups are presented for the rating of social studies and science in Table 2.

It was observed that at the Grade 5 level only one of the differences between means (usefulness) was significant. All of the differences, except one, were in the hypothesized direction. At the grade 8 level, all of the differences were in the expected direction, and all but one of the six differences were significant at the $p<.01$ level.

TABLE 2

COMPARISON OF MEAN RATINGS BY SUBCULTURE STUDENTS AND MAIN CULTURE STUDENTS FOR SOCIAL STUDIES AND SCIENCE

	Dimension	Subculture			Main Culture			Difference	Level of Sig. of Diff.
		N_1	$\bar{X}_1$	S_1	N_2	$\bar{X}_2$	S_2	$\bar{X}_1 - \bar{X}_2$	
Soc. St. Gr.5									
	Evaluation	43	.66	4.79	164	1.32	4.29	- .66	-
	Usefulness	43	2.23	4.73	164	3.69	4.43	-1.46	-
	Difficulty	43	-.11	3.92	164	.01	3.74	- .12	-
Sci. Gr. 5									
	Evaluation	40	.15	5.57	159	1.24	5.15	-1.09	-
	Usefulness	40	.98	4.70	159	3.03	4.58	-2.05	< .05
	Difficulty	40	1.50	4.45	159	.75	4.27	.75	-
Soc. St. Gr.8									
	Evaluation	75	-1.39	3.62	270	1.27	3.78	-2.66	< .01
	Usefulness	75	2.05	4.47	270	4.06	3.89	-2.01	< .01
	Difficulty	75	- .56	3.73	270	- .13	3.16	- .43	-
Sci. Gr. 8									
	Evaluation	76	-1.49	4.76	268	1.13	4.74	-2.62	< .01
	Usefulness	76	.75	4.75	268	3.75	4.36	-3.00	< .01
	Difficulty	76	-2.93	3.77	268	-1.06	3.80	-1.87	< .01

These results were considered to be in reasonable agreement with expectations. It was thought likely that, as the subculture students progressed through school, social studies in particular, and science would be perceived as having diminishing relevance to the subculture, and that this perception would be reflected in their responses on the instrument. The fact that the results bore out this hypothesis was taken as evidence of validity of the instrument.

Sex differences were examined as another criterion for validity. Of the two subjects compared, social studies and science, teacher and administrator opinion was that girls, expecially older girls, would have less favourable attitudes toward science than would boys.

TABLE 3

COMPARISON OF RATINGS OF MALES AND FEMALES FOR SCIENCE

			Evaluation		Usefulness		Difficulty	
Group	Grade	N	Mean	S.D.	Mean	S.D.	Mean	S.D.
Males	5	97	1.49	5.10	2.49	4.81	.91	3.98
Females	5	105	.55	5.31	2.60	4.57	.80	4.59
Significance of Difference			N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
Males	8	187	1.77	4.23	4.04	3.86	-1.14	3.69
Females	8	158	-.87	5.16	1.97	5.15	-1.87	4.04
Significance of Difference			$p < .01$	$p < .01$	$p < .01$	$p < .01$	N.S.	N.S.

Table 3 sets forth the results for science. It can be observed that at the Grade 5 level the differences in mean scores for evaluation, usefulness, and difficulty were relatively small, not statistically significant, and not consistently in any direction. However, at the Grade 8 level, two of the three comparisons (for evaluation and for usefulness) demonstrated differences of mean scores in the direction of less favourable attitudes by girls, of a magnitude which was statistically significant at the $p<.01$ level.

The results for social studies, not presented here, revealed for each dimension at each grade level that differences in mean scores were not statistically significant at the $p<.05$ level.

The sex differences found were not great in the lower grades. For older students some differences were found; these were in the expected direction.

It was possible to obtain an estimate of achievement from the final marks for Grade 4 and Grade 7 as awarded in the preceding June. The attitude measures were taken in the following autumn, starting in September. For one school, results of a Lorge Thorndike intelligence test (raw scores) were available for Grade 8 students. The correlations between achievement or mental ability and attitude toward school subjects are presented in Table 4.

TABLE 4

ATTITUDE-ACHIEVEMENT AND ATTITUDE-INTELLIGENCE

Science Attitude Factor		Science Achievement			Intelligence		
		Males	Females	Total	Males	Females	Total
	N=	239	266	511	70	60	132
Evaluation		0.010	.050	-.001	.101	-.197	-.052
Usefulness		.117	.161*	.116**	.179	-.119	-.013
Difficulty		-.045	.127*	.048	-.049	-.003	-.029

* $p < .05$

** $p < .01$

Of 18 possible comparisons three were statistically significant, two at the 0.05 level and one at the 0.01 level. It was concluded that for these students, attitudes toward school subjects were relatively independent of either achievement or intelligence.

Finally, a factor analysis involving the Stage 4 data was carried out. All of the data were combined; that is, ratings of all school subjects investigated for both Grades 5 and 8 served as the basis for a single factor analysis. After rotation to the varimax criterion the three factors that resulted were found to be identical to those found in the earlier stages.

Stage 5

During the administration of the measure in Stage 4, Clarke had observed that some students did not know the meanings of some words, usually one of a pair. He had explained the meanings at the time. It was decided that for such words, a simplified and brief definition would be provided. Although the factors previously found had remained intact, the test-retest reliability was not as high as would be desirable, particularly that based on a one-month time lapse. School examinations and report cards had been scheduled during the interval; teachers and administrators expected this might affect students' attitudes toward school subjects. It was decided that further reliability studies should be tried. In this connection, it was decided to lengthen the usefulness and difficulty scales so that each would consist of eight bipolar pairs, and to reduce the evaluation scale, which had the highest reliability, to eight. These procedures would maintain the length of the instrument at the desired 24 bipolar pairs. It was also deemed desirable to try out the scale at the high school (Grade 11) level.

The first task was to revise the scale. Three bipolar pairs of words with the lowest factor loadings were eliminated from the evaluation scale. These were still-active, sad-happy, and alert-dull. Two bipolar pairs of words were added to the usefulness scale (impractical-practical, and harmful-advantageous) while one

(smart-stupid) was dropped because of adverse teacher reaction and because it loaded about equally on both the evaluation and usefulness scale. Two bipolar pairs were added to the difficulty scale: complicated-simple, and undemanding-rigorous.

Based on the experience in administering the scale during Stage 4, the instructions to students were revised, particularly by adding additional examples of the rating procedure used. For the five-point rating scale (very much, a bit, neither, a bit, very much) the revised instructions pointed out that the rating "neither" could be used if the student felt that the bipolar pair of words did not apply to the subject being rated, or if it did apply but the rating was midway along the scale.

TABLE 5

RELIABILITY ESTIMATES BASED ON FOUR TESTING OCCASIONS: ATTITUDE TOWARDS MATHEMATICS

Factor		Occasion 2	3	4
Evaluation	1	.76	.80	.73
	2		.83	.84
	3			.88
Usefulness	1	.63	.82	.65
	2		.67	.74
	3			.78
Difficulty	1	.63	.86	.70
	2		.75	.85
	3			.80

Occasion code:

1 = day 1, 1st test
2 = day 1, 2nd test (30 min. after 1st test)
3 = day 2, 1st test (1 week after day 1)
4 = day 2, 2nd test (30 min. after 1st test)

correlations between occasions 1 and 2, and 3 and 4 are estimates of reliability that are not subject to trait unreliability. The correlation between occasion 1 and 3 appears alone in Table 3.

The revised instrument was again administered by Clarke in seven schools. In four of the schools, one week later the measure was again administered to the same students for the same subjects. At the Grade 11 level, in one large school, the school subjects rated were Mathematics, Social Studies, Science, Language, Literature, Physical Education, then Mathematics again, all at one sitting. This was possible because of the traditionally longer class period in the high school. In this school, one week later, the same students again rated the same subjects.

The test-retest reliability, based on a 30-minute and on a one-week time lapse for 191 Grade 11 students rating the school subject mathematics is shown in Table 5.

Thus, for the school subject mathematics, the test-retest reliability for the evaluation factor with a 30-minute time lapse was 0.76 and 0.88. There were four possible estimates of the test-retest reliability after a one-week time lapse: day 1 first test vs. day 2 first test, day 1 first test vs. day 2 second test, day 1 second test vs. day 2 first test, and day 1 second test vs. day 2 second test. These were found to be 0.80, 0.73, 0.83, and 0.84 respectively. With a 30-minute interval, the estimate of reliability should not be subject to trait unreliability.

TABLE 6

TEST-RETEST RELIABILITY BASED ON A ONE-WEEK TIME LAPSE

Subject	Scale	Reliability Gr. 5 (N = 85)	Gr. 8 (N = 292)	Gr. 11 (N = 191)
Social Studies	Evaluation	.64	.75	.74
	Usefulness	.63	.67	.67
	Difficulty	.61	.76	.71
Science	Evaluation	.73	.75	.79
	Usefulness	.65	.70	.72
	Difficulty	.73	.70	.78
Phys. Ed.	Evaluation	.67	.76	.77
	Usefulness	.78	.68	.74
	Difficulty	.78	.70	.72
Mathematics	Evaluation	.75	.80	.80
	Usefulness	.75	.79	.82
	Difficulty	.76	.81	.86
English Lang.	Evaluation			.69
	Usefulness			.71
	Difficulty			.73
English Lit.	Evaluation			.70
	Usefulness			.72
	Difficulty			.70
All Subjects	Evaluation	.72	.79	.84
	Usefulness	.82	.77	.84
	Difficulty	.79	.77	.82

Test-retest reliabilities based on a one-week time for Grades 5, 8 and 11 for a number of school subjects are shown in Table 6. It will be noted that in general, these reliabilities range from 0.65 to 0.85 for the various subjects and dimensions. In general, the reliabilities are highest for Grade 11 students and lowest for Grade 5 students. The persisting vocabulary difficulty may account in part for this finding.

Each scale had been altered somewhat so another factor analysis was performed to determine whether the original factors remained intact. Table 7 shows the results.

TABLE 7

VARIMAX ROTATED FACTOR LOADINGS BY GRADE

		Evaluation			Usefulness			Difficulty		
		5	8	11	5	8	11	5	8	11
+	nice-awful	.72	.73	.77						
-	boring-interesting	.73	.77	.80						
-	unpleasant-pleasant	.69	.71	.80						
-	dislike-like	.72	.71	.78						
+	bright-dull	.75	.78	.83						
-	dead-alive	.75	.79	.82						
+	lively-listless	.68	.76	.78						
+	exciting-tiresome	.67	.72	.73						
-	useless-useful				.66	.75	.78			
+	important-unimportant				.63	.83	.85			
-	impractical-practical				.65	.78	.82			
-	worthless-valuable				.71	.84	.87			
+	helpful-unhelpful				.65	.82	.85			
-	unnecessary-necessary				.74	.85	.84			
-	harmful-advantageous				.55	.67	.72			
+	meaningful-meaningless				.64	.72	.75			
-	hard-easy							.64	.72	.74
+	light-heavy							.61	.60	.72
+	clear-confusing							.65	.70	.72
-	complicated-simple							.66	.78	.82
+	elementary-advanced							.72	.45	.59
-	strange-familiar							.45	.55	.57
+	understandable-puzzling							.58	.65	.67
+	undemanding-rigorous							.27	.41	.55

Percentage variance accounted for:
Grade 5 - 53%
Grade 8 - 63%
Grade 11- 68%

It can be observed that the factors remained intact, the loadings were similar for all three grades and were consistent, that is, each of the word pairs had its highest loading on the same three factors for all three grades. One of the loadings for Grade 5 was low (.27) but the bipolar word pair was retained because the loadings for Grade 8 and 11 were satisfactory, and because it was hard to find good pairs for the difficulty factor.

Table 7 presents the bipolar pairs of words in their final form for each of the three factors. Also included is the scoring key on the +2, +1, 0, -1, -2 scale. Thus, a student rating of a subject as "very much" nice would be assigned a score of +2. If a student rated a subject "a bit" boring it would score -1, and so on.

The words which were defined, with the simplified definition, were as follows:

listless (inactive, lazy)
tiresome (makes a person feel tired)
advantageous (brings good or gain)
heavy (a lot of work)
confusing (mixes a person up)
advanced (beyond the beginning level)
puzzling (hard to understand)
rigorous (has to be exactly right).

Discussion

The School Subjects Attitude Scales is a reliable and valid instrument for measuring attitudes toward school subjects. Reliability varies by subject, being higher for mathematics and science than for social studies and English. It varies by grade, being higher for the higher grades. It varies by dimension, but with no clearly discernible pattern. In general, the reliabilities found for each dimension ranged from 0.65 to 0.85.

Validity was determined prior to the final revision of the scale. However, since it was ascertained that the factors remained unchanged and the reliability for each increased, it was assumed that the findings would apply to the final revision. For each criterion used (expert opinion, student preferences, type of school, cultural differences, and sex differences) student ratings were in the expected direction and differences were frequently of considerable magnitude. At the Grade 5 level, cultural differences and sex differences were small, and for social studies, sex differences were small.

Experience in administering the scale supports the view that faking is a real problem. Measures such as anonymity of results, administration by a neutral person, and assurances that results cannot and will not reflect in any way on the student, may reduce faking.

Three dimensions which appeared to be very stable were found in students' attitude toward school subjects. They were identified as *evaluation* (e.g., boring-interesting), *usefulness* (e.g., worthless-valuable) and *difficulty* (e.g., complicated-simple).

The semantic differential type instrument developed had the versatility desired, but was hampered somewhat by the difficulty in finding bipolar pairs of words which could be easily comprehended by students, particularly those in lower grades. Despite this, however, the instrument seems to be of potential use in grades as low as the fifth.

The instrument in its present form might be useful to teachers and to school systems. However, a few limitations should be borne in mind if use is contemplated. First, it is important to note that in the developmental stages of the instrument the administrator was perceived by the students to be an outsider who could legitimately claim to have no power over the students and who stated that individual responses would be scored by a machine and not revealed to anyone. A second limitation is that no norms have yet been developed, therefore the instrument can be used to make local comparisons only.

Further development of the School Subjects Attitudes Scales will likely be undertaken. The major task will be the construction of norms. It is expected that separate norms will be needed for different subjects, different grades, for boys and girls, and perhaps for different types of schools. In addition, separate norms for time of year may be required, e.g., September, December, February, and May. (As this article goes to press, the authors report that they have commenced work on developing norms for Alberta.)

References

Maguire, T.O. Semantic differential methodology for the structuring of attitudes. *American Educational Research Journal*, 1973, *10*(4), Fall, 295-306.

Nyberg, V. R. & Clarke, S.C.T. *Technical report on the School Subjects Attitudes Scales.* Edmonton, Alberta: Alberta Education, May, 1979.

Osgood, C. E., Suci, G., & Tannenbaum, P. H. *The measurement of meaning.* Urbana, Illinois: University of Illinois Press, 1957.

The Alberta Journal of Educational Research *Vol. XXVIII, No. 2, June, 1982, 188-192*

BOOK REVIEWS

THE CHARACTERISTICS AND NEEDS OF ADULTS IN POSTSECONDARY EDUCATION. *By Lewis C. Solmon and Joanne J. Gordon.* Toronto: Lexington Books, 1981.

If you're an administrator concerned with continuing education, enrollment levels, or adult education and need to plan for the future of adult enrollment in your postsecondary institution, then you'll probably find Solmon and Gordon's book of interest. The book is a compendium of statistics about adult students (people over 21) and contains more statistics than you probably want to read (86 tables worth of data).

While the data reviewed in the book are based on a somewhat biased sample of adults in the United States and may be a bit out of date (sample years 1969-1978), the information should be useful in a heuristic sense for (i) making predictions about the future of adult education, and (ii) helping postsecondary institutions develop schemes to attract more adult students.

One tricky area in planning for adult education is the definition one chooses for adult. The first chapter of the book considers this dilemma but does not reach a definitive solution. Solmon and Gordon discuss the ramifications of different definitions and indicate how different definitions have come about. They do draw the reader's attention to the important differences between looking at adult students and at adults in education and have chosen to look primarily at adult students. The definition Solmon and Gordon utilize was chosen because the data base was a population of students all of whom were over 21, entering postsecondary education for the first time, attending school full time, and attending an institution that returned enough questionnaires to be considered valid. Solmon and Gordon's population therefore includes only those adult students in some sort of a degree program. While not all of the adults were seeking a degree, the program they were in would allow it. While Solmon and Gordon go beyond their data base's restrictive definition and consider all adult students surveyed (including some part time students and those attending institutions without a high return rate), the sample of adults surveyed, while large in number, may not be representative of the adult student population. Solmon and Gordon, however, note this in their book and, where appropriate, indicate the possibility of bias in their sample. They also reviewed additional information based on other samples and thus extended the above definition of adult student. These extensions increase the utilizability of the book and increase its heuristic guidance. In general, one could conclude that information which applies to over-21 full time students also applies to the general adult student population, but takes on even broader perspectives when applied to the expanded population.

Considering the sample of adult students used and the expanded "market" that really exists, the book provides one with enough information to decide what types of adult students your institution may be interested in attracting and what types of programs and assistance your institution needs to provide in order to attract these students.

The general assessment of adults who are presently attending postsecondary institutions appears to be that adults usually attend smaller institutions which provide a broad range of programs which either are directly related to specific career goals, or are of current interest to the adults in increasing their general level of knowledge and are probably associated with business, education, health, or social science. The message to universities and four-year colleges contained in these statistics seems to be "if you would like to attract more adult students, then you need to expand your typical types of program offerings." This suggestion, however, should not be interpreted to mean that universities and four-year colleges should attempt to do what community and two-year colleges are doing. Nor should it be interpreted to mean that universities and four-year colleges should "lower" their standards. What is implied is that universities and four-year colleges should attempt to do what they are presently doing, but should attempt to "package" their programs in different ways and should be more flexible in the ways they can accommodate the needs of adult students.

The wide variety of adults attending postsecondary institutions and the wide range of demographics attached to these adults (see Chapters 1 & 2) imply that there are more than enough adults to go around. More and more women are returning to school and the adult population in school is growing older (greater numbers of adults in their 30's, 40's, and 50's). If the different types of postsecondary institutions simply continue to do what they are presently doing (with respect to types of programs offered) and expand the way they offer their programs (more flexible transfer credit, more flexible time considerations for both the offering of courses and completing courses, and more flexibility in the ways students meet program requirements), all institutions should be able to share the adult market.

Chapters 1 & 2 provide the demographic statistics referred to above and should be consulted for any specifics in which one is interested. Chapters 3-7 provide the background as to what types of institutions adults typically choose to attend, and the reasons for those choices. The information in these chapters should be particularly useful in helping institutions design alternative packages of their normal program offerings so that the new packages are more "attractive" to adult students. One factor of importance to institutions in creating new "packages" is that adult students are more interested in the quality and reputation of the "package" than they are in the overall reputation of the institution. Another important factor for adults is the specialty services offered by the institution. These services range all the way from direct financial support to day care facilities. Institutions need to be aware of the greater difficulties adults have when coming back to school because of demands of families and jobs.

One important area of service adults will need surrounds the type of background skills and knowledge they bring with them to learning situations. Institutions will either need to provide some type of remedial program or will need to be innovative in the ways they establish instructional situations for the adults. The creation of innovative practices appears to be the more "attractive" to adults because it allows for greater flexibility and it takes into account their greater maturity and experience. This appears to be one of the major reasons community and two-year colleges have enjoyed such a large share of the adult student "market."

The final chapter of the book (Life Goals) notes that adults do not fit into neat categories. About the only general conclusion one can draw is that adults appear to be less concerned with "making it" in their profession. They are more interested in a breadth of perspectives and in "daily competence" in their lives. The current group of adult students exhibits interests in being able to do things well and in helping others do things well, a fact consonant with the realization that today's adults are/were the traditional students of the 1960's.

The overall conclusion I've drawn from the book is that if postsecondary institutions take into account the greater specificity of adult needs, the broader perspective that adults bring with them because of their greater maturity, the need for teaching procedures that vary from "traditional teaching," the need for more variety and flexibility in support services (day care, remedial help, financial help, study skills, etc.), the need for flexibility in the time of offering and/or the time to complete courses, and in the type of course evaluations, then postsecondary institutiions will attract greater numbers of adult students.

R. Schultz
Department of Educational Psychology
The University of Alberta

INTELLIGENCE, LEARNING, AND ACTION. *By Richard R. Skemp.* Toronto: John Wiley & Sons, 1979.

In this book Richard Skemp presents a new conception of intelligence and, more importantly, the function of intelligence. Intelligence is reflected in goal-directed behavior that functions to enhance the prospects of survival. Readers who admired his previous book, *The Psychology of Learning Mathematics*, will appreciate the continuance of certain themes from that book to this latest offering. Skemp has not done a flip-flop and gone off in some totally different direction. Although they appear in different garb, many of his important previous contributions such as reflective intelligence and primary and secondary concepts are still there to be taken away and appreciated. Certainly Skemp has succeeded in revealing the evolution of his thinking about intelligence.

Though bridges can be built from his previous work to this latest book, I would not want to underplay the potential value and usefulness of his latest innovative insights and theories about human intelligence. I am hesitant to try to summarize these insights and theories because Skemp ranges quite widely in this book, but, at the risk of oversimplification, I will try to hit what appear to be the high points. In a fundamental sense, Skemp says that survival, evolution, and learning are intimately related. In this model, intelligence is important because the nature of it is a certain ability to learn. Then the function of intelligence becomes the development of director systems that are able to adapt to environmental conditions so that our chances of survival might be enhanced. For Skemp, director systems seem to be a necessary construct in order to see behavior as being goal-directed. Of course defining behavior as goal-directed is quite necessary when survival is the sine qua non.

Much of the first part of the book is devoted to an elaboration of the nature and function of intelligence as it relates to various aspects of being human. For example, emotions become important sources of information about progress relative to goal states and antigoal states. For Skemp, this relative progress is determined by information received through perception that fuses our inner reality and actuality, producing modifications in both as a result. Out of these dynamics, knowledge is built as structures of inner reality that have withstood the testing against actuality. These structures of inner reality are hypothetical constructs that are revealed in the functioning of director systems.

In Skemp's model, director systems are an important component in defining intelligent learning. Skemp conceptualizes learning as a process which involves two levels of director systems in a manner very reminiscent of his previous notion of reflective intelligence. Skemp posits a delta-one director system whose state can be changed in various ways, among them through teaching, and a delta-two director system that has the delta-one system as its operand and whose primary function it is to take the delta-one system toward states of optimal functioning. What are most interesting and important for education are the interactions between director systems, e.g., consider the interaction between the director systems of teacher and student.

Among the many remaining constructs and ideas that Skemp uses in his model there is one more that deserves special mention and that is the notion of resonance. Skemp uses resonance as a mechanism for the explanation of the workings of human memory and also as an explanatory construct in such activities as concept learning. Indeed, resonance is an important notion in Skemp's model.

As noted previously, this book is wide ranging as it looks at intelligent learning with a cybernetic behavior-as-goal-directed point of view. In fact, the scope of this book has caused some unease for me in my efforts to summarize it. It is hoped that the reader can at least get some flavor of the content of this book even if glaring, though inadvertent, omissions were quite likely made.

In general, I think this book makes an important contribution to educational literature. Skemp should be lauded for taking some of the latest theories of cognitive science and interpreting them for the nonspecialist and, also, drawing out some of the implications of these theories for uses in educational settings. These are difficult tasks and he has done them well.

Though I like the book and recommend it to the reader, there are some aspects of the proposed model that I think are questionable. In fairness, Skemp admits to the speculative nature of parts of the model and I certainly agree with him in his statements about the importance of theories and models, right or wrong, in guiding our reflective thought about intelligence and learning. Nevertheless, there seem to be some parts of the model which are philosophically questionable. To begin, defining behavior as goal-directed is troublesome when one tries to go deeper than to say behavior is brought about for some superficial purpose. Skemp seems to tie behavioral goal-directedness to the primary need for survival. That seems problematic to me because of the obvious large chasm between survival needs and behavioral goal-directedness. From another view, I think I can still ask who or what directs the director systems. Another philosophical troublespot becomes evident in distinguishing between inner reality and actuality. That distinction is not as easy to maintain as common sense would have us believe. As Hegel noted, that distinction can make sense only if a third party observer is included in the inquiry. If I have an inner reality and there is an actuality out there, then how am I going to recognize it? A final nagging problem for me, this one not necessarily philosophical, concerned the role of emotions in

the model. Skemp seems to see them as information-bearers about status or progress vis-à-vis goal states or antigoal states. I had a profound sense of unease when the rich world of emotion was reduced to a finite set of roles relating to goal states. Perhaps that was not Skemp's intent so that the reader should assume only that a small subset of emotions were being discussed. Even conceding that, I felt the role of emotions was overdetermined in Skemp's account, belying the complexity of that part of humanness.

Notwithstanding the critical comments made above, I felt that there were other parts of the book that could be very helpful to people who need to understand the workings of human intelligence, particularly as it operates in educational settings. For those people, I would particularly recommend a careful reading of chapters 14-16. Skemp was careful in relating the content of these chapters to his model of learning, but I almost had a feeling that it wasn't necessary to buy his theory entirely to get much of value from these chapters. They are written by a man who has thought long and carefully about many of the perennial problems facing education. They propose solutions which are not piece-meal or superficial, but they are also not hopelessly romantic and out of touch with reality. In short, these chapters will give a reader a sense of hope that there are possibilities for improvement in education if we are committed to improvement and are willing to learn intelligently about what we are doing in education.

Alton T. Olson
Department of Secondary Education
The University of Alberta

Research Grants Information

Social Sciences and Humanities Research Council of Canada

RESEARCH GRANTS PROGRAM

Revisions to the Research Time Stipend to include independent scholars

The Social Sciences and Humanities Research Council of Canada wishes to announce that as of May 15, 1982 (the next deadline for the receipt of research grant applications), the guidelines governing the provision of a *research time stipend* within the framework of the Research Grants Program will be extended to include all scholars in the Canadian research community, whether or not they hold a full-time teaching position in a postsecondary institution. For those scholars who were not previously eligible to apply for a research time stipend, the provisions outlined in the Research Time Stipend guidelines in the *1981 Guide for Applicants* will apply, subject to the following modifications:

1. Two categories of research time award will be available to research grants applicants who do not hold a full-time teaching position at a postsecondary teaching institution. The established maximum amounts of these awards are $21,400 and $35,000 per annum. The latter category is reserved for senior scholars who have already demonstrated a substantial commitment to the Canadian research community through significant contributions to scholarship.
2. The total maximum amount of research time which may be claimed by these applicants will be twelve months in any three-year period. Subject to this limit, there will be no restriction on the amount of time which may be taken in any one year.
3. Employed scholars and those who become employed during the tenure of a grant will be required to furnish evidence that their employers have agreed to release them for the period of time proposed. They must also demonstrate that their proposed research is not directly related to their employment and that it will not interfere with their normal duties.
4. All research proposals which include a stipend request are evaluated by the assessors and Council adjudication committees according to the usual assessment criteria of the Research Grants Program. In addition, the assessors and committees will be asked to recommend to Council both the appropriate stipend category for each applicant and the exact amount of research time that should be allowed. Recommendations will be based on the qualifications of the applicant and the particular needs of the project for which support has been requested.

The Research Time Stipend guidelines as outlined in the *1981 Guide for Applicants* will continue to apply to all scholars who hold full-time teaching positions in universities and other postsecondary institutions. Current Council regulations with respect to the eligibility of applicants and of research projects will also continue to be applied in all cases.

For further information, please contact:

Research Grants Division
Social Sciences and Humanities Research Council of Canada
P.O. Box 1610
Ottawa, Ontario
K1P 6G4

February 1982

(français au verso)

Conseil de recherches en sciences humaines du Canada

PROGRAMME DES SUBVENTIONS DE RECHERCHE

Allocation de dégagement pour la recherche offerte aussi aux chercheurs autonomes

The Conseil de recherches en sciences humaines offre normalement une *allocation de dégagement pour la recherche* dans le cadre des subventions de recherche. Il désire annoncer par la présente qu'à compter du 15 mai 1982 (la prochaine date limite pour les demandes de subventions), il élargira l'offre de cette allocation à toute la communauté de chercheurs canadiens, qu'ils soient ou non affectés à une tâche d'enseignement postsecondaire à plein temps. Les directives contenues dans le *Guide des candidats 1981* concernant cette allocation s'appliqueront donc dorénavant à ceux qui jusqu'ici n'étaient pas admissibles, parce que non universitaires. Les modifications suivantes devront cependant être retenues, dans leur cas.

1. Pour les chercheurs autonomes (non affectés à l'enseignement postsecondaire à plein temps) il y aura deux catégories d'allocation, allant respectivement à un maximum de $21,400 et de $35,000. Cette deuxième catégorie sera réservée à des chercheurs chevronnés qui se seront déjà révélés profondément engagés dans la communauté des chercheurs canadiens et qui auront déjà fourni une importante contribution à la connaissance et à l'érudition.
2. Les candidats sollicitant une allocation pourront compter sur un maximum de douze mois sur trois ans, sans limite particulière de temps sur chaque année.
3. Les chercheurs qui détiennent un emploi, ou en obtiennent un pendant la validité de leur subvention de recherche, doivent démontrer que leur employeur consent à les libérer pour la période de temps correspondant à l'allocation demandée. Ils doivent aussi démontrer que leur projet de recherche n'est pas directement relié à leur emploi et qu'il n'y aura pas conflit avec leurs tâches normales.
4. Tous les projets de recherche qui comportent une allocation de dégagement sont soumis à l'évaluation des pairs et aux comités d'adjudication du Conseil, selon les critères normaux du programme de subventions de recherche. De plus, ces deux instances d'appréciation devront aviser le Conseil tant sur la catégorie d'allocation que sur le montant précis auquel le Conseil devrait s'engager. Une telle recommandation devra reposer sur la compétence du chercheur et sur les besoins particuliers du projet soumis.

Il est à remarquer que le *Guide des candidats 1981* continue à fournir les directives régulières en ce qui concerne l'allocation de dégagement offerte aux chercheurs qui enseignent à plein temps dans les universités ou autres institutions postsecondaires. Les règlements actuels qui gouvernent l'admissibilité des candidats et des projets de recherche demeurent inchangés.

Pour tout renseignement supplémentaire, on doit s'adresser à:

La Division des subventions de recherche
Conseil de recherches en sciences humaines
C.P. 1610
Ottawa, Ontario
K1P 6G4

Février 1982

PREPARATION OF MANUSCRIPTS

1. All manuscripts must be typewritten, double spaced, and submitted in duplicate. An abstract of approximately 100 words in length, typed on a separate page, should be provided.
2. Tables must be numbered in Arabic numerals with the word 'Table' centered and in capital letters, e.g., TABLE 4. The heading of the table is to be centered below and typed in capitals. The format of tables should conform to the specifications in the APA Publications Manual.
3. Graphs and charts should be used only if essential. They must be carefully prepared on separate sheets in India ink, ready for reproduction. Graphs must be properly labelled using Arabic numerals, e.g., Figure 3.
4. Each table or figure should be presented on a separate page. The position of tables and graphs should be clearly indicated within the text by inserting at the relevent point the phrase (Insert Table 2 here).
5. References should appear in parentheses following the reference citing the author's name (unless the name appears in the text), the year of publication, and page number if appropriate. For direct quotations, the reference should be cited and the page number given in brackets before the final punctuation of the quotation. The references should be listed alphabetically by authors' last names at the end of the manuscript under the heading, *References*.
6. Explanatory notes, numbered consecutively and identified in the text with a superscript, may be included under the heading of *Notes*. They should be double spaced and placed at the end of the manuscript immediately preceding the *References*. The citing of references and quotations in the *Notes* should conform to the procedures outlined in No. 5 above.
7. Spelling shall conform to the *Oxford English Dictionary* or to the *Dictionary of Canadian English,* except in the case of direct quotations which must conform to the original. Editorial alterations will be made if necessary.
8. In matters of style, the APA Publications Manual is considered definitive.

PRINTED BY
THE UNIVERSITY OF ALBERTA
PRINTING DEPARTMENT